Journal of
Experimental Psychology

VOL. 25, NO. 3 SEPTEMBER, 1939

RELATIONS BETWEEN EXPERIMENTALLY INDUCED MUSCULAR TENSION AND MEMORIZATION *

BY FREDERICK A. COURTS
University of Missouri

Of the numerous investigations having to do with mental work, many have led to the discovery of psychological and physiological factors which affect the efficiency of performance in various situations. One such investigation is that of Bills (1), in which it was found that muscular tension induced by gripping the handles of two dynamometers brought about a significant improvement in performance in various mental tasks. While some of the investigations subsequent to Bills' original work have tended to substantiate his results, others have failed to do so. These studies have been reviewed by Bills (2).

The work of Freeman (8), and that of Block (4), suggests that there may be a functional relationship between the amount of experimentally induced muscular tension and the level of performance, in which tensions below or above the optimum for a given performance are detrimental, as compared with the effect of optimal tension. This suggestion is compatible with results reported by Jacobson (9), who has shown that in a condition of extreme relaxation mental activity occurs on only a low level, if at all. Stauffacher (11)

* This investigation was performed in the Psychological Laboratory of Stanford University. The writer is indebted to Dr. E. R. Hilgard and Dr. C. P. Stone for their advice and assistance during the course of the experiment.

has recently attacked this problem of the functional relationship between the amount of experimentally induced muscular tension and the level of performance in memorizing nonsense syllables. To illustrate the relationship he has presented an hypothetical curve based on memorization scores obtained under four degrees of muscular tension. Although all three of the experimentally induced tensions which he used resulted in an improvement of performance over that observed under conditions of normal muscular tension, Stauffacher has extrapolated the curve to indicate the decrement in performance which might be expected from the results of Freeman (**8**) and Jacobson (**10**) to occur under extreme degrees of tension and relaxation. Stauffacher's data fail to show an actual decrement from normal performance under any of the conditions which he used. Although his prediction that extreme degrees of muscular tension will result in such a decrement appears reasonable, it appears desirable that this relationship be worked out in greater detail. One of the purposes of the present investigation is to do this.

The existence of individual differences in susceptibility to the facilitative influence of muscular tension on mental work was noted by Bills in his original investigation (**1**, p. 249). Later, Block (**4**) reported that her subjects tended toward optima at different points in the series of tensions which she used. More recently Bills and Stauffacher (**3**), and Stauffacher (**11**) have further confirmed the existence of such differences, suggesting that poor performers are more apt to benefit from experimentally induced tension than are good ones, the latter being adversely affected. Neither of these authors, however, reports data adequate to substantiate the alleged differential effect of tension on good and poor performers.

As an explanation of the inadequately demonstrated differential influence of muscular tension for good and poor performers, Stauffacher has proposed the hypothesis that ". . . good performers were already working at a very high level of efficiency (correlated with a certain degree of muscular tonicity) and that the increased tonus placed their level of muscular tonus over the facilitating degree and therefore

caused a decreased efficiency of performance" (11, p. 44). Although Stauffacher does not explicitly formulate it, the presumption is that the beneficial influence of experimentally induced muscular tension in the case of poor performers would be explained on the assumption that they already are working under a degree of tension below the optimum. Thus, the addition of experimentally induced tension would produce a resultant tension which is nearer the optimum. Stauffacher's hypothesis, although unnecessary until the differential effect of tension for good and poor performers is adequately demonstrated, is a very attractive one. It assumes, however, that the amount of tension over and above that normally maintained by the subject during memorization bears a linear relationship to the physical force necessary to maintain various weights in suspension or to maintain various amounts of pressure on a dynamometer. In other words, the possible interaction of two types of activity (mental and physical effort) in determining total bodily tension has been ignored. It is possible that the differential influence of experimentally induced tension is a function of the resultant effect produced by these two factors.

The first objective of the present investigation is to work out in detail the relationship between the level of performance in memorization and various degrees of experimentally induced muscular tension. The second objective is to ascertain the importance of certain factors which may play a part in determining individual differences in susceptibility to the facilitative influence of such tension. Three of these factors will be considered: (1) the relative level of memorization under normal conditions, without additional tension; (2) the amount of muscular tension normally accompanying the memorization process itself; (3) the combined tension produced by memorization and physical effort occurring together.

Apparatus

Muscular tension was measured in terms of the amplitude of the knee-jerk as indicated by the amount of quadriceps-thickening in response to the tendon-tap. This method of measuring tension has been discussed in detail elsewhere (6).

The material to be memorized consisted of 6-unit lists of 3-place consonant (nonsense) syllables made up of items found by Witmer (14) to have an association

value of less than 60%. The syllables were presented by means of a standard memory drum, manufactured by C. W. Perrollaz, of the Department of Psychology of Cornell University. The memory drum was modified by the addition of a 10-watt electric light, to illuminate the material and a rotary switch which was connected in series with a small light. This light served to inform the experimenter of the turning of the drum. The switch also operated a relay which released the stimulus hammer, synchronizing the stimulation of the knee-jerk with the presentation of the syllables. The tendon-tap, when tension was measured during memorization, thus occurred at a constant interval following the appearance of an item in the window of the memory drum. This interval, as measured by a Cenco impulse counter, was 0.23 second. The

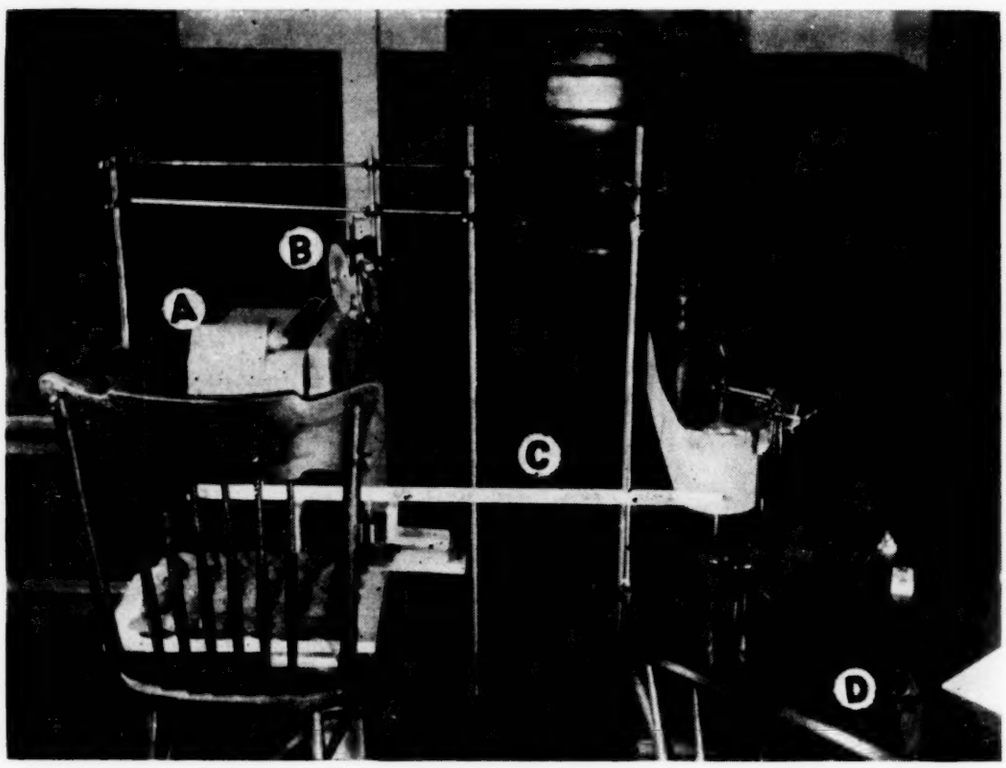

FIG. 1. Apparatus. *A*. Memory drum. *B*. Dynamometer. *C*. Recording lever. *D*. Key for releasing hammer.

work of Bowditch and Warren (5) suggests that variability in this relationship would be a source of variability of the knee-jerk.[1]

Muscular tension was induced by squeezing a Smedley hand-dynamometer. This device was equipped with a small contact of spring brass, which was connected with a 3-volt flashlight lamp situated in front of the subject. The contact could be set on the dial of the dynamometer so that when the indicator hand reached a given point the light would be turned on. The flexibility of the contact was such as to permit a range of 3 kilograms, within which the light would remain on. The handles of the dynamometer were wrapped with strips of sponge-rubber, which were covered with cloth.

The entire apparatus, with the screens removed, is shown in Figure 1.

[1] Although the activity preceding the tendon-tap in this experiment is quite unlike that studied by Bowditch and Warren, it is interesting to note that their results indicate only a very slight facilitation of the knee-jerk when the motor response precedes the tendon tap by an interval of 0.23 second.

Experimental Procedure

The plan of the experiment required that each subject participate on three separate days. Subjects were scheduled so that the three sessions came at the same hour on each of three days within a week. In no case did more than two days elapse between two successive experimental sessions. The 60 subjects were male college students, enrolled in a class in elementary psychology.

First day: The first experimental session required approximately 35 minutes. After 30 measurements of the amplitude of the knee-jerk had been recorded, the subject was asked to squeeze as hard as possible on the dynamometer handles for a period of 30 seconds. The reading at the end of that time was taken as a base for the various degrees of tension to be used throughout the experiment. Ten records of the knee-jerk were secured under each of 6 degrees of dynamometer tension [2] equal to various fractions of the basic amount as determined in this way. The tensions occurred in the following order: ½, 0, ⅜, ⅛, ¾, ¼, ¼, ¾, ⅛, ⅜, 0, ½. Five records of the amplitude of the knee-jerk were obtained during each presentation of each of the degrees of tension. The remainder of the first experimental session was devoted to preliminary practice in memorization. This consisted of 5 trials on each of 4 lists of syllables. Each item was exposed for 3 seconds and there was an interval of 18 seconds between trials.[3] The subjects were given the following typewritten instructions before the practice in memorization:

> The work which you will do during this experiment will be the memorization of lists of consonant syllables. (A consonant syllable consists of three consonants, such as BMT.)
>
> The exposure device in front of you will present a series of consonant syllables. You should give your undivided attention to the syllables as they appear in the window. After we have gone through the list once, there will be a short interval of rest, then we shall go through the list again. This time, and in all succeeding trials, you must attempt to *spell* each of the syllables before it appears. That is, when the X's at the beginning of the list appear in the window, you should attempt to spell the first syllable in the list *before* it appears; then when that syllable appears, try to spell the next syllable, etc., on through the entire list.
>
> Always try to spell the syllable in the list which comes *after* the one at which you are looking.
>
> Always check on the accuracy of your attempted anticipation by noticing the syllable which turns into the window, so that you may correct your errors.

---

[2] The term 'experimentally induced tension' has been used in other investigations to refer to the condition resulting from muscular effort. This expression is too inclusive, since other aspects of the experimental situation, such as the very presence of the experimenter, may bring about a considerable change in the condition of tension manifested by the subject. In the following discussion the term 'dynamometer tension' will be used to refer to changes resulting directly from dynamometer effort, regardless of whether these changes are neural, or muscular, or both. Similarly the term 'memorization tension' will be used to refer to those changes which accompany memorization.

[3] This relatively long rest period was introduced in order to lessen fatigue resulting from maintaining dynamometer tension during memorization in the third experimental session.

Second day: The second session required about 30 minutes. Thirty measurements of the amplitude of the knee-jerk were obtained under no dynamometer tension. The dynamometer performance at the end of 30 seconds of maximum effort was again determined. The subject then memorized 4 lists of syllables, following the procedure described above. However, during each trial for 2 of the lists, 3 knee-jerks were recorded. The stimuli were distributed at random throughout each trial. The lists were presented in the same order for all subjects, but with half of the subjects the knee-jerk was stimulated during lists 1 and 4. With the other half it was stimulated during lists 2 and 3.

Third day: This session of the experiment required about 60 minutes. The procedure for 30 of the subjects included the preliminary dynamometer test, as on the preceding days, and the memorization of 12 lists of syllables under the various degrees of dynamometer tension which had been used on the first day. Here the subject was instructed to squeeze the dynamometer handles hard enough to turn on the light when the X's appeared in the window of the exposure device. The light was to be kept on until the first blank space appeared at the end of the list. During the 18-second interval between presentations of the list, the subject was allowed to rest. There was an interval of 60 seconds between lists.

The 12 lists used on this day were presented in 3 different orders. Ten subjects followed each. The orders are shown in Table 1.

TABLE 1

ORDERS OF PRESENTATION

Order of Lists	Dynamometer Tension											
	½	0	⅜	⅛	¾	¼	¼	¾	⅛	⅜	0	½
A	1	2	3	4	5	6	7	8	9	10	11	12
B	6	5	4	3	2	1	12	11	10	9	8	7
C	3	4	1	2	6	5	8	7	11	12	9	10

The procedure for the remaining 30 subjects was the same, except that 30 records of the knee-jerk were obtained at the beginning of the session, and one knee-jerk was recorded during each presentation of each list of syllables. The tendon-tap occurred at random, accompanying the presentation of a different syllable in each trial.

RESULTS

The influence of dynamometer tension on the amplitude of the knee-jerk observed during the first day of the experiment has been reported elsewhere in a discussion of the knee-jerk as a measure of tension (**6**). In view of this fact, the results of the first experimental session will not be discussed here.

Reliability of the Memory Scores.—Memory scores were recorded in terms of the total number of letters correctly anticipated (and correctly placed within the syllable) during

the 4 recitation trials on each list of syllables. In order to determine the reliability of the scores obtained under each degree of dynamometer tension, the sums of the letters correctly anticipated in trials 1 and 3 of the first list and in trials 2 and 4 of the second list were correlated with the sums obtained in trials 2 and 4 of the first list and trials 1 and 3 of the second list. These coefficients are presented in Table 2.

Reliability of the Dynamometer-Test.—The correlation between the dynamometer reading after 30 seconds of maximum effort on the first day and that obtained on the second day is $r = .79$. The correlation between the first and third days is

TABLE 2

RELIABILITY OF MEMORY SCORES UNDER VARIOUS TENSIONS ($N = 60$)

Dynamometer Tension	r	r (Corrected by Spearman-Brown formula)
0	.76	.86
1/8	.80	.89
1/4	.90	.95
3/8	.85	.92
1/2	.81	.90
3/4	.87	.93

$r = .81$. Although this measure may be affected by factors other than the strength of the subject, it is fairly reliable.

Influence of the Tendon-Tap on Memorization.—Introduction of the knee-jerk as a measure of memorization tension is not without complications, for there is the possibility that the tendon tap, occurring as it does somewhat irregularly and without warning, may adversely affect the memorization. On the second day of the experiment, memorization scores were obtained from each subject under normal conditions, and also while stimulating the knee-jerk 3 times during each presentation of the list of syllables. The mean score obtained on 2 lists of syllables under each of these conditions is shown in Table 3. The C. R. of 3.6 indicates that the tendon-tap did adversely affect memorization during the second experimental session. However, when the results of the two subgroups of 30 subjects are compared for the third experimental session, it is apparent that no such effect occurred on that day.

TABLE 3

INFLUENCE OF THE TENDON-TAP ON MEMORIZATION ($N = 60$)

Conditions	Mean Memory Score	σ	σm
Control	60.7	18.4	2.4
With Knee-Jerk	53.4	17.1	2.2

M Difference	7.3
* σ M Difference	2.0
C. R.	3.6

* $\sigma m\,(x - y) = \dfrac{\sigma\,(x - y)}{\sqrt{N}}$.

This comparison is shown in Table 4. A possible explanation of the absence of interference from the stimulation of the knee-jerk during the third session is that the tendon-tap occurred only once during each presentation of the list of syllables, whereas in the preceding session it had occurred 3 times during each presentation. Another possible explanation is that the subjects may have become adapted to the

TABLE 4

INFLUENCE OF THE TENDON-TAP ON THE RELATIONSHIP BETWEEN MEMORIZATION AND DYNAMOMETER TENSION ($N = 30$ IN EACH GROUP)

	Dynamometer Tension					
	0	⅛	¼	⅜	½	¾
Mean Score without K.-J.	57.0	59.6	66.6	63.6	61.8	51.0
With K.-J.	53.1	60.1	62.3	58.5	58.8	51.1
Difference	3.9	−0.5	4.3	5.1	3.0	−0.1
* σ Diff.	4.2	4.4	5.2	4.9	4.5	5.4
C. R.	.93	−.11	.83	1.04	.67	−.02

* σ diff. $= \sqrt{\sigma^2 m_1 + \sigma^2 m_2}$.

distracting influence of the tendon-tap. Whatever the true explanation may be, the failure to observe a detrimental effect resulting from the stimulation of the knee-jerk during the third session provides justification for pooling the data obtained during this session from the two sub-groups of 30 subjects in order to obtain a more representative picture of the influence of dynamometer tension on memorization.

Memory Performance Under Various Degrees of Dynamometer Tension.—The means of the memory scores under various conditions of tension are presented in Table 5. This relationship is presented graphically in Fig. 2. The significance of

TABLE 5

MEAN MEMORY SCORES UNDER VARIOUS DEGREES OF DYNAMOMETER TENSION ($N = 60$)

	Dynamometer Tension					
	0	⅛	¼	⅜	½	¾
Mean Score	55.1	59.8	64.4	61.0	60.3	51.0
σ	16.2	17.0	20.6	19.3	17.4	20.8
σ_m	2.1	2.2	2.7	2.5	2.2	2.7

the increase or decrease under each condition of dynamometer tension is indicated by the critical ratios in Table 6.

It is apparent that the facilitative influence of dynamometer tension on memorization is functionally related to the amount of tension under which the memory performance occurs. The optimal amount of tension is that induced by

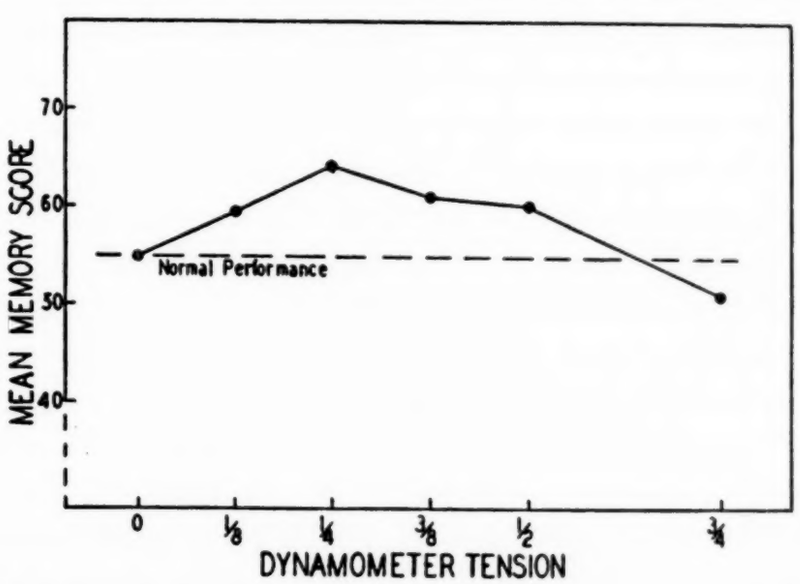

FIG. 2. Mean memory scores under various degrees of dynamometer tension ($N = 60$).

dynamometer pressure equal to ¼ of the performance at the end of 30 seconds of maximum effort. Dynamometer tension of ⅜ also results in a significant gain in memory scores. The gain under ½ dynamometer tension, while not as great,

TABLE 6

SIGNIFICANCE OF THE INCREASE IN MEMORY SCORES UNDER VARIOUS DEGREES OF DYNAMOMETER TENSION ($N = 60$)

Conditions	Diff.	r	* σ Diff.	C.R.
1/8 minus 0.....	4.3	.57	2.0	2.2
1/4 minus 0.....	9.3	.58	2.2	4.2
3/8 minus 0.....	5.9	.67	1.9	3.1
1/2 minus 0.....	5.2	.56	2.0	2.6
3/4 minus 0.....	−4.1	.63	2.1	−2.0

* σ diff. $= \sqrt{\sigma^2 m_1 + \sigma^2 m_2 - 2r_{12}\sigma m_1 \sigma m_2}$.

approaches significance. Performance under 3/4 dynamometer tension is not significantly below normal. It is, however, significantly below the best performance under dynamometer tension. That the curve would continue to fall with higher degrees of tension than those used here is suggested by the progressive decline with successive degrees of tension above the optimum.

It is interesting to note here that Stauffacher (11), taking the dynamometer performance at the end of 60 seconds of maximum effort as a base, found that dynamometer tension equal to 1/2 of that amount resulted in the best performance. In the present experiment, the reading at the end of 30 seconds of maximum effort was taken as the base. Consequently the determination of the subjects' strength in the present experiment would be higher in comparison with their actual strength than it was in Stauffacher's experiment. Thus, our finding that the optimal tension is 1/4 of the base does not conflict with his observation that it is 1/2 of the base, as determined in his experiment.

The more rapid rise and decline shown by Stauffacher's curve may be due to the fact that he used only 4 experimentally determined values in plotting the relationship. In the present investigation, 6 conditions of tension were used. Furthermore, the successive increases in tension were in smaller intervals in the region of maximal facilitation of memorization. Thus, our representation of the function is in greater detail than that presented by Stauffacher. In general, however, results of the present experiment substantiate his

finding that there is a functional relationship between the amount of experimentally induced muscular tension and its influence on memorization.

The Influence of Dynamometer Tension on Good and Poor Memorizers.—In order to compare amounts of change in performance scores representing different levels of ability, it is necessary to express those scores in terms of equal units. If this is not done, a change of a given number of units in the raw score may represent a change in the level of performance which is quite different for scores taken from different levels in the range of abilities. Bills and Stauffacher (3) have presented their evidence for the differential effect of muscular tension on good and poor performers in terms of raw time and error scores, without demonstrating that those scores are in equal units of ability. Thus, although tension may have influenced good and poor performers in different ways, their treatment of the data provides neither an accurate measurement nor a conclusive demonstration of this alleged phenomenon.

The memory scores obtained in the present investigation were tabulated into 6 overlapping distributions corresponding to the 6 degrees of tension used. Scale values were then computed for the various obtained scores by the method proposed by Thurstone (12) (13), and applied to practice data by Woodrow (15). The relationship between the scaled values and the corresponding raw scores is shown in Figure 3. The relationship is practically linear, which indicates that the raw scores obtained in this experiment may be used in comparing the effect of dynamometer tension on good and poor performers. Because of their convenience the raw memory scores will be used in the following analysis. (A detailed description of the procedure followed has been presented in an unpublished paper (7).)

A group of 21 poor performers was obtained by taking the 7 subjects who achieved the lowest scores under conditions of no dynamometer tension for each of the 3 orders of presentation of the lists of the syllables on the third day. Similarly, a group of 21 good performers was selected by taking the 7

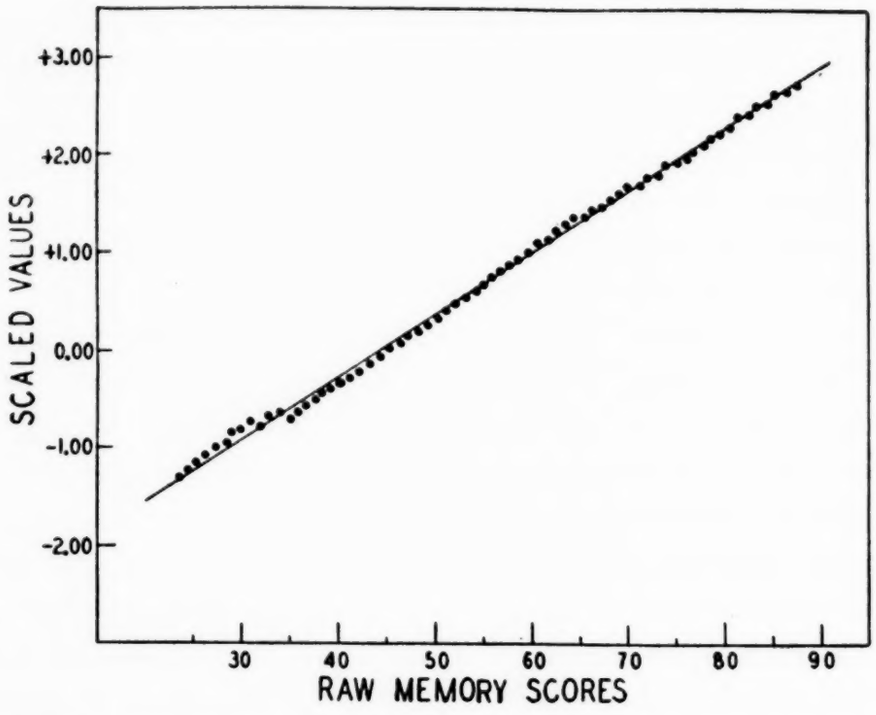

FIG. 3. Relationship of the scaled values to corresponding raw memory scores

best scores for each order. The mean memory score obtained on two lists by the good memorizers was 71.0. That of the poor memorizers was 40.1. The difference of 30.9 is 8.6 times its standard error.

The means of the memory scores obtained by the two groups under the various tensions are represented in Figure 4. This figure is based on the values shown in Table 7. The optimal degree of dynamometer tension for both groups is that induced by pressure equal to ¼ of the base. The values in Table 8 indicate that the good memorizers gained significantly only under ¼ dynamometer tension, whereas poor memorizers gained significantly under tensions of ⅛, ¼,

TABLE 7

MEAN MEMORY SCORES OBTAINED BY GOOD AND POOR PERFORMERS UNDER VARIOUS DEGREES OF DYNAMOMETER TENSION (N = 21 IN EACH GROUP)

Group	Dynamometer Tension					
	0	⅛	¼	⅜	½	¾
Good Performers...	71.0	70.2	80.0	76.0	71.4	62.3
Poor Performers....	40.1	49.9	52.1	50.2	47.0	40.0

and ⅜. The loss shown by good memorizers under ¾ dynamometer tension approaches significance, while poor memorizers show no loss under any of the degrees of dynamometer tension.

On the basis of Stauffacher's hypothesis we should expect that the optimum of dynamometer tension would occur earlier in the series of tensions for good memorizers than it would for poor ones. When the optimum dynamometer tension for each subject was expressed in units of ⅛ of the

TABLE 8

SIGNIFICANCE OF GAINS OF GOOD AND POOR MEMORIZERS UNDER DYNAMOMETER TENSION

Good Memorizers

Dynamometer Tension	Mean Gain	σ	σm	C.R.
⅛	−0.8	15.5	3.4	−0.2
¼	8.9	12.9	2.9	3.1
⅜	5.0	13.8	3.1	1.6
½	0.4	16.5	3.7	0.1
¾	−8.7	14.9	3.3	−2.6
		Poor Memorizers		
⅛	9.8	14.5	3.2	3.1
¼	12.0	16.5	3.7	3.2
⅜	10.0	14.2	3.2	3.1
½	6.9	16.3	3.6	1.9
¾	−0.2	12.3	2.7	−0.1

base, it was found that the average for good performers was 2.7. The mean of the optima for poor performers was 3.0. The standard error of this difference is 0.4, indicating that the optimal dynamometer tension is the same for good and poor memorizers. Moreover, Figure 4 shows that the average performance of the poor memorizers improves more rapidly with successive degrees of tension than does that of the good performers. If, as Stauffacher suggests, good performers normally work under a condition of tension near the optimum, additional tension would be expected to facilitate their performance early in the series of tensions. Figure 4 does not substantiate this expectation. However, the more rapid

decline of the curve representing the performance of good memorizers is in line with Stauffacher's suggestion.

Since the curve for good performers in Fig. 4 declines slightly between the points representing performance under 0 and 1/8 dynamometer tension, whereas the curve for poor performers rises between these points, it may be inferred that the two groups are not really different in ability. If the group of good performers included those subjects who by chance had obtained high scores during the trials on which the differentia-

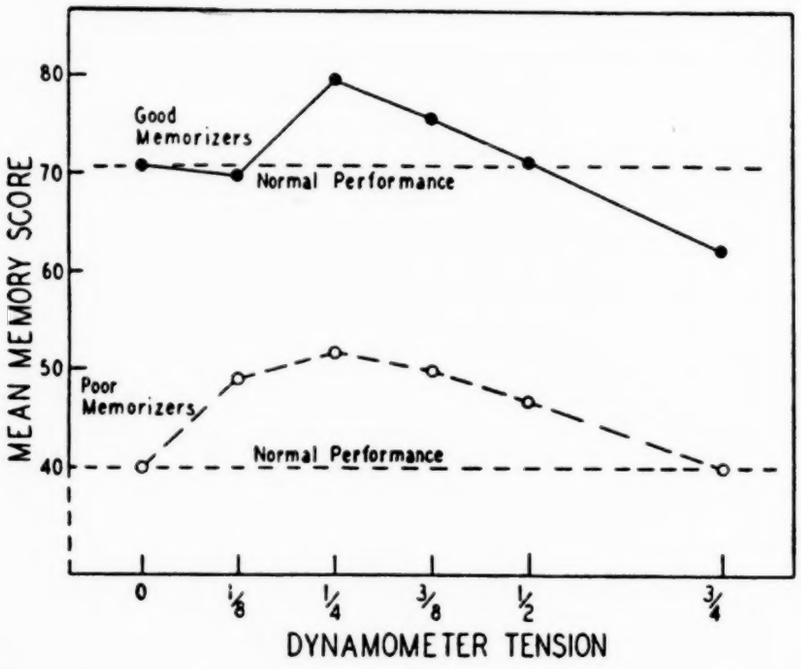

FIG. 4. The influence of dynamometer tension on good and poor performers (N = 21 in each group).

tions was based, the point on the curve representing those trials would be spuriously high. Likewise, the performance of the poor group would have been lower than their true ability, if the differentiation were on the basis of chance fluctuations. If that were true, the poor group would be more apt in other trials to be lower and the good group higher in relation to the chance-determined points of reference. The two groups, however, are different in ability, as is shown by a comparison of their performance on the second day of the experiment. The average score of the good memorizers on 4 lists of syllables on that day was 137.3. The average of the poor memorizers

was 93.5. The difference of 43.8 is 5.1 times its standard error.

The values presented in Table 9 indicate that none of the differences between the gains of the two groups is statistically significant.

TABLE 9

SIGNIFICANCE OF DIFFERENCES BETWEEN THE GAINS OF GOOD AND POOR MEMORIZERS UNDER DYNAMOMETER TENSION (N = 21 IN EACH GROUP)

	Dynamometer Tension				
	⅛	¼	⅜	½	¾
Diff.	10.6	3.1	5.0	6.5	8.5
$\sigma_{Diff.}$	4.7	4.7	4.4	5.2	4.3
C. R.	2.2	0.7	1.1	1.2	2.0

The number of individual subjects in each group whose memory scores show a gain or loss from performance under normal conditions of tension is shown for each degree of dynamometer tension in Table 10. It is apparent that dynamometer tension affects neither good nor poor performers consistently.

The data of the present investigation show no significant tendency for poor performers to benefit more than good ones as a result of dynamometer tension. Nevertheless, Stauff-

TABLE 10

FREQUENCY OF INCREASE AND DECREASE IN MEMORY SCORE UNDER CONDITIONS OF DYNAMOMETER TENSION FOR GROUPS OF GOOD AND POOR MEMORIZERS (N = 21 IN EACH GROUP)

Poor Memorizers

	Dynamometer Tension				
	⅛	¼	⅜	½	¾
Gain	16	17	15	11	10
Loss	5	4	6	9	10
No Effect	0	0	0	1	1
Good Memorizers					
Gain	11	18	13	8	8
Loss	10	3	8	12	13
No Effect	0	0	0	1	0

acher's suggestion that the degree of tension normally accompanying memorization may be an important factor in determining the differential influence of dynamometer tension may yet be acceptable, since the normal level of memorization tension may not be highly correlated with memory performance.

The Differential Influence of Dynamometer Tension as a Function of the Tension Normally Present During Memorization.—To determine the amount of tension present during memorization, the amplitude of the knee-jerk was recorded while the subject was memorizing two lists of consonant syllables on the second day of the experiment. The mean amplitude of the 30 responses obtained in this way for each subject was compared with the amplitude of response obtained under normal conditions on the same day. The mean amplitude of the last 25 of the 30 responses recorded at the beginning of the second experimental session was taken as the amplitude of the 'normal' knee-jerk. This was done because the first 5 responses showed a very rapid decrease in amplitude, while the decrement during the last 25 responses was very slight. The progressive decrease in amplitude with continued stimulation of the knee-jerk on the second day of the experiment is shown in Fig. 5. The mean increase in the amplitude of the knee-jerk during memorization was found to be 10.5 mm.[4] The S. D. of the increases was 12.5, indicating the existence of wide individual differences in memorization tension.

In order to obtain two groups of subjects differing in the amount of tension exhibited during memorization, the 7 subjects showing the greatest and the 7 showing the least increase in the knee-jerk during memorization were selected for each order of presentation of the lists of syllables on the third day. The 21 subjects in the least increase group manifested a mean increase in amplitude of 3.7 mm. The mean increase for the most increase group was 22.2 mm. The difference

[4] The amplitude of the knee-jerk is reported in terms of the deflection of the recording stylus. In the present experiment this is approximately 4 times the amount of vertical movement of the surface of the leg directly over the belly of the quadriceps muscle.

of 18.5 mm is 8.8 times its standard error. Table 11 gives the mean memorization scores obtained by the two groups under the various degrees of dynamometer tension during the third day of the experiment. The largest differences between the groups are to be found under ⅛ and ¼ dynamometer tension. The difference of 7.2 under ⅛ dynamometer tension is 1.5 times its standard error. The difference of 6.9 under ¼ tension is 1.3 times its standard error. The slight, and statistically unreliable, differences between these two

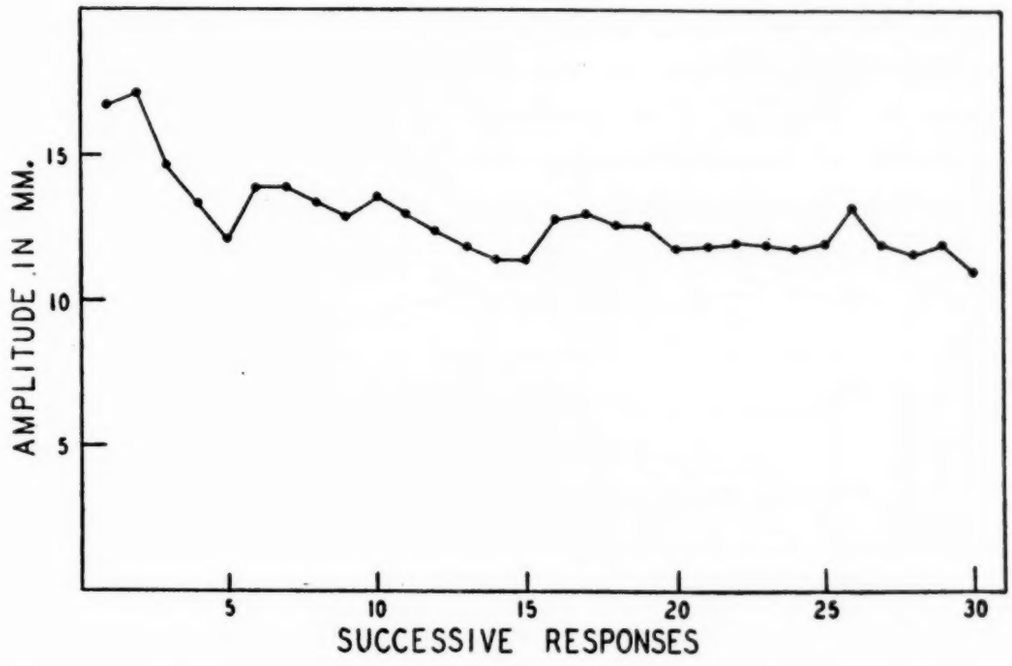

FIG. 5. Progressive decrement in the amplitude of the knee-jerk during the second experimental session (N = 60).

groups in susceptibility to the influence of dynamometer tension on memorization suggest that the amount of tension normally present during memorization is not an important factor in determining the influence of additional muscular tension.

The Differential Influence of Dynamometer Tension as a Function of the Total Tension Produced by the Combined Influence of Mental and Muscular Effort.—During the third experimental session one knee-jerk was elicited from each of 30 subjects during each presentation of each list of syllables. Thus 10 responses were recorded for each subject during

TABLE 11

MEAN MEMORY SCORES OBTAINED BY THE TWO GROUPS UNDER VARIOUS DEGREES OF DYNAMOMETER TENSION (N = 21 IN EACH GROUP)

Group	Dynamometer Tension					
	0	⅛	¼	⅜	½	¾
Least Increase.....	54.3	62.8	66.6	59.7	58.1	49.7
Most Increase.....	55.3	55.0	59.7	60.8	59.0	53.3

memorization under each condition of tension. The influence of the combination of memorization and various degrees of dynamometer tension on the amplitude of the knee-jerk is shown in Fig. 6. For purposes of comparison, the mean amplitude of response under the various dynamometer tensions alone is shown for the same 30 subjects. Each of the

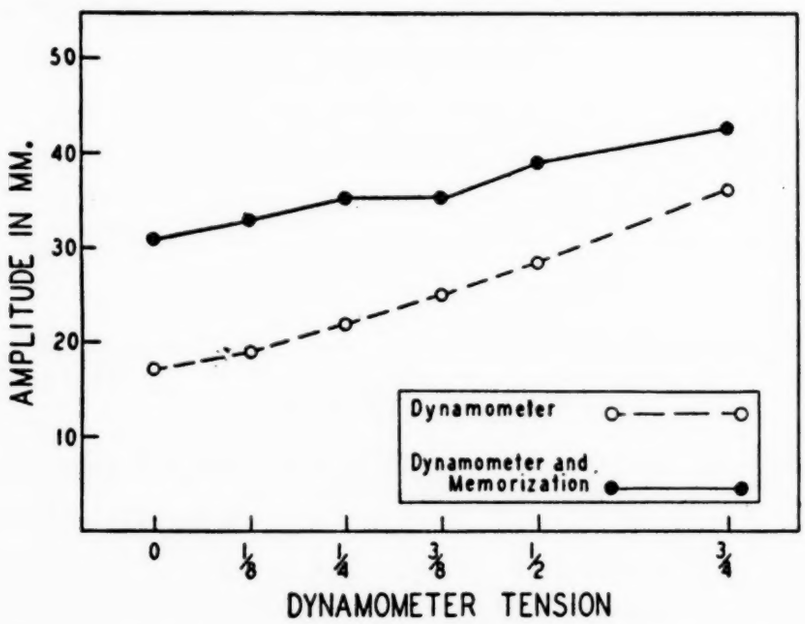

FIG. 6. Changes in the amplitude of the knee-jerk resulting from the combined influence of mental and muscular effort (N = 30).

points on the two curves represents the mean of the averages of 10 responses obtained from each of the 30 subjects. The mean amplitudes of response during memorization under dynamometer tension have been increased by 8.2 mm, since this was the average decrement of the normal knee-jerk during the three days of the experiment for the 30 subjects whose results appear in Fig. 6. Thus, the comparison of the two

curves is not complicated by the day-to-day changes in the normal amplitude of response during the experiment.

Since it was desired to determine the relationship of the differential influence of dynamometer tension on memorization to the total tension resulting from the effects of mental and muscular effort occurring together, a group of 10 subjects who showed the greatest average gain in memorization under the various dynamometer tensions was selected. The mean of the average gains for this group was 17.2. The mean of

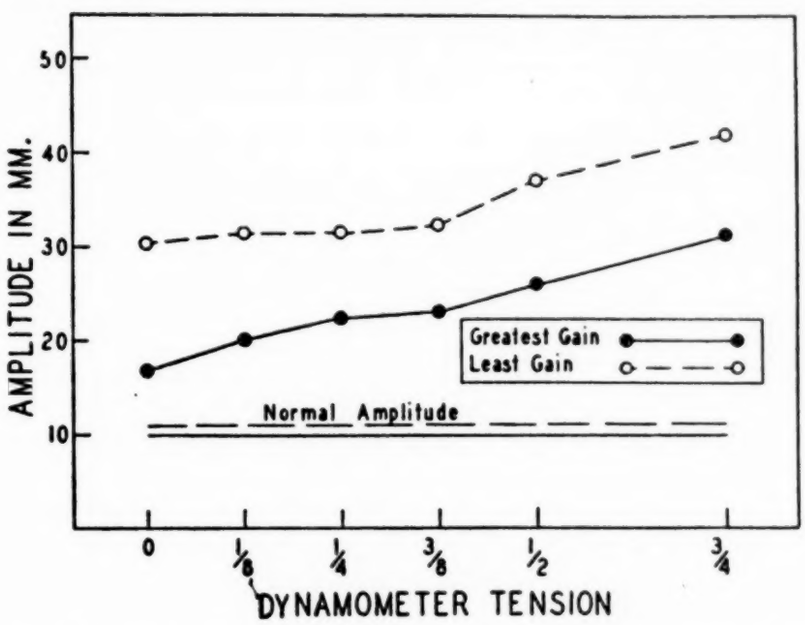

FIG. 7. Changes in the amplitude of the knee-jerk during memorization under conditions of dynamometer tension for subjects showing the greatest and the least gain in memory scores as a result of such tension (N = 10 in each group).

the gains for the 10 subjects who showed the least gain under dynamometer tension was −7.3. The t-value for this difference of 24.5 points in average gain is 6.77, which indicates that the probability of such a difference arising by chance is less than .01. The groups of 10 are conclusively different in susceptibility to the influence of dynamometer tension on memorization.

Figure 7 represents the changes in the amplitude of the knee-jerk during memorization under the various dynamometer tensions for the two groups of subjects. The base-line of 'normal' amplitude for the two groups, as determined on the same day, is also shown in Figure 7. The 10 subjects who

show the greatest average gain in memorization under dynamometer tension show a relatively slight increase in tension resulting from memorization alone. The subjects who fail to benefit greatly from dynamometer tension show a greater increase in tension during memorization. This tends to confirm Stauffacher's suggestion that in the case of those subjects who already work under a high degree of tonus an additional amount of experimentally induced tonus will result in a condition which is beyond the optimum and thus cause a decrease in the efficiency of performance. It will be recalled,

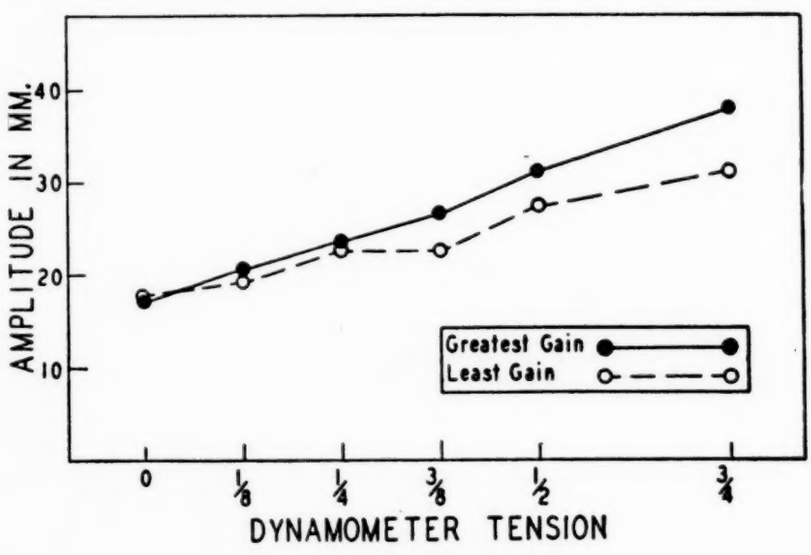

FIG. 8. Changes in the amplitude of the knee-jerk with various dynamometer tensions for subject showing the greatest and the least gain in memory scores as a result of such tension (N = 10 in each group).

however, that when two groups of subjects were selected on the basis of the amount of increase in tension accompanying memorization no significant difference was found in the influence of additional dynamometer tension. Although it is possible that the normal level of memorization tension may have a part in determining the influence of dynamometer tension on memorization, the results of the present investigation suggest that such an influence, if it exists, is slight.

The present analysis (Figure 7) shows that the combination of dynamometer tension and memorization effort results in greater tension for the 10 subjects who fail to benefit from dynamometer tension, and in less tension in the case of

subjects who memorize more efficiently under dynamometer tension. That this is not the result of differences between the groups in respect to the change in the amplitude of the knee-jerk resulting from dynamometer tension alone is shown by Figure 8. Here the two groups are compared in terms of the influence of dynamometer tension alone on the knee-jerk. The slight difference in the influence of dynamometer tension on the amplitude of the knee-jerk for the two groups is in the opposite direction from the differences resulting from the combined influence of mental and muscular effort, shown in Figure 7. These data provided evidence that the individuals whose memory performances benefit most from dynamometer tension are the ones who exhibit less tension as a result of the combination of memorization and dynamometer effort.

Conclusions

Stauffacher's (11) statement of the relationship between the level of memorization and the amount of experimentally induced muscular tension has been substantiated by the results of the present investigation. With successive degrees of dynamometer tension, induced by gripping with the right hand, memorization is progressively more efficient until the optimal tension is reached. The optimum in the present experiment was found to be the amount of tension produced by dynamometer pressure equal to ¼ of the reading at the end of 30 seconds of maximum effort in a preliminary trial. Higher degrees of tension result in successive decrements in performance. Under tension induced by pressure equal to ¾ of the base, memorization falls below the normal level of performance.

The results of the present investigation show no evidence for a differential effect of experimentally induced tension for good and poor memorizers. Group results show that both good and poor memorizers perform best under ¼ dynamometer tension. The means of the optimal tensions for good and poor performers are not significantly different. The number of subjects among the good and poor memorizers who show either a gain or a loss in performance under tension

is approximately the same for each of the 6 degrees of tension used.

The influence of dynamometer tension on memorization is not significantly different for subjects who normally maintain a high degree of tension during memorization and subjects who normally memorize under a low degree of tension.

Subjects who show little or no facilitation of memorization under dynamometer tension show greater tension resulting from the combination of dynamometer tension and memorization effort, whereas subjects whose memorization is greatly improved by the addition of dynamometer tension show less tension resulting from the synthesis of the two tension producing factors.

(Manuscript received March 15, 1939)

References

1. Bills, A. G., The influence of muscular tension on the efficiency of mental work, *Amer. J. Psychol.*, 1927, **38**, 227–251.
2. Bills, A. G., Facilitation and inhibition in mental work, *Psychol Bull.*, 1937, **34**, 286–309.
3. Bills, A. G., and Stauffacher, J. C., The influence of voluntarily induced tension on rational problem solving, *J. Psychol.*, 1937, **4**, 261–271.
4. Block, H., The influence of muscular exertion upon mental performance, *Arch. Psychol.*, 1936, No. 202, Pp. 49.
5. Bowditch, H. P., and Warren, J. W., The knee-jerk and its physiological modifications, *J. Physiol.*, 1890, **11**, 25–64.
6. Courts, F. A., The knee-jerk as a measure of muscular tension, *J. Exp. Psychol.*, 1939, **24**, 520–529.
7. Courts, F. A., *Relations between experimentally induced muscular tension and memorization.* Unpublished Doctor's Thesis, Stanford University, 1938.
8. Freeman, G. L., The facilitative and inhibitive effects of muscular tension upon performance, *Amer. J. Psychol.*, 1933, **45**, 17–52.
9. Jacobson, E., *Progressive relaxation.* Chicago: Univ. Chicago Press, 1929.
10. Jacobson, E., Electrical measurement of neuromuscular states during mental activities, *Amer. J. Physiol.*, 1930, **91**, 567–608; **94**, 22–34; **95**, 694–702, 703–712; **97**, 200–209.
11. Stauffacher, J. C., The effect of induced muscular tension upon various phases of the learning process, *J. Exp. Psychol.*, 1937, **21**, 26–46.
12. Thurstone, L. L., A method of scaling psychological and educational tests, *J. Educ. Pychol.*, 1925, **16**, 433–451.
13. Thurstone, L. L., The unit of measurement in educational scales, *J. Educ. Psychol.*, 1927, **18**, 505–524.
14. Witmer, L. R., The association value of three-place consonant syllables, *J. Genet. Psychol.*, 1935, **47**, 337–360.
15. Woodrow, H. H., The scaling of practice data, *Psychometrika*, 1937, **2**, 237–247.

THE SPEED OF MONOCULAR ACCOMMODATION*

BY RICHARD H. LEUKART
Ohio State University

From time to time the speed of accommodation has been investigated with increasing refinement of technique. The earliest investigators (Volkmann (**13**), Vierordt (**12**)) counted the number of fixations in a given time that could be made to a near and a distant point. Later investigators (Aeby (**1**), Barrett (**4**), Seashore (**10**)) introduced the finger reaction time as a method of recording the time of a single change in fixation.

The best solution to the problem of timing the reaction appears to be the objective regulation of the exposure time of the crucial object. It is this approach, with regard to apparatus and method, which Ferree and Rand (**7**) have taken. Their apparatus consists of a series of cut-out discs, revolved by an unbalanced pendulum, so arranged that a near test object is seen, followed by a far one (or vice versa) and the exposure times of each can be controlled and varied. Thus, if the exposure time of the first is well above threshold and the subject is not able to designate the position of the capital E of the second exposure, it may be inferred that the time of exposure of the second object was shorter than the 'adjustment time' of the subject.

These investigators have wisely named this movement an 'adjustment' rather than 'accommodation' in the strict ophthalmological sense. The accommodation movement involves only the interrelations of ciliary muscle, zonule, and lens which alter the refractive power of the lens. The 'adjustment' necessary with this apparatus involves convergence (since binocular observation was used) and change of lateral fixation, or line of sight (because the far object was seen to one side of the position in which the near object had been or was

* This investigation was conducted at Ohio State University under the direction of Professor F. C. Dockeray.

to be), as well as accommodation to the different distances of the objects from the eyes of the observer.

Tefft and Stark (11), working at Mitchell Field, Long Island, used the Ferree and Rand apparatus in an attempt to demonstrate speed of 'accommodation' as a practicable test for fliers. The confusion of accommodation with adjustment is unfortunate. One of the conclusions of the Tefft and Stark study is that slow accommodation, with the possible exception of cases in which presbyopia has begun, is associated with other visual defects which are more easily detected. This may be true of slow 'adjustment,' in which all of the visual apparatus is involved, but is probably not true of accommodation in the strictest sense.

The most recent work on this problem is reported in a paper by Banister and Pollock (3). They measured far to near accommodation by using two fall-tachistoscopes. The large (far) letters, exposed at a distance of 610 cm from the subject were E T P S V R G. The small letters, 41.5 cm away and above the level of the far letters, were C G O D Q W K N M H. The ability to name the near letter after a momentary exposure was the criterion of accommodation.

Average accommodation times for three subjects were .62″, .52″, and .40″. These times include that required to change the fixation, move the eye upward, as well as accommodation time. When the vertical angular distance of the two test objects was increased from 1°5′ to 2°32′ and both test letters were approximately the same distance from the eye, the time of .62″ was reduced to .32″.

There are several important defects, from the standpoint of measuring the accommodation time alone, in the methods reviewed above. The reaction technique is probably unreliable; at best it is less convenient and sensitive than the method depending upon the objective control of the length of exposure of the test objects. The necessity of changing the fixation before or during the time when accommodation takes place introduces the action of the extra-ocular muscles, which action is not a part of accommodation. This is true also of binocular determinations, since convergence movements then

accompany accommodation. Finally, the test objects used have not been entirely satisfactory or, in some cases, have not been used to best advantage. The present investigation was designed to eliminate these factors in the measurement of accommodation time.

Problem

This investigation was planned in an attempt to devise a method to measure the time required for the accommodative mechanism of the eye to effect distinct vision. The requirements of this apparatus were that it should not necessitate a change of fixation while the accommodative movement was taking place, and that the time of exposure of the second object could be controlled so that the ability to name this object could be used as a criterion of the completion of the accommodation movement.

Method

A Meumann Time Sense apparatus was used as an important part of this device. This apparatus, was placed upon the table with its axis in the horizontal position, thus resting on two of the three feet and upon the edge of the rim. Only one contact, the star wheel, was used and this was fastened securely at the top of the rim, which was now in a vertical plane. On the star wheel of this contact was mounted a very light cardboard disc, with a diameter slightly larger than that of the wheel, in such a way that the disc revolved with the wheel. With the apparatus in operation, the disc was revolved through 60 degrees at each contact of either of the two revolving arms of the apparatus.

Upon the cardboard disc were mounted two small rectangular mirrors, about 8 mm by 10 mm, in positions corresponding to two adjacent arms of the star wheel. These mirrors were placed at slightly different angles with regard to the observer's eye so that one reflected the near object and the other the distant object to the same point. When the apparatus was in motion, one revolving arm moved one mirror into position, exposing one test object. As the other arm revolved the star wheel through 60 degrees, the second mirror was moved into exposure position, thus reflecting the second test object immediately after and in the same position as the first object. The first contact arm again moved the star wheel through 60 degrees, thus removing the second object from view. An illuminated white field was visible when neither mirror was in the exposure position. The angle between the two revolving arms may be varied to control the time of exposure of the second object.

Since letters of the alphabet were used as test objects, it was necessary to introduce a second mirror. The first mirror (on the star wheel) reversed the letters while the second mirror reversed the reflection, thus presenting them as though viewed directly. This second mirror was fastened to the edge of the table just to the right of the ob-

server's head and was placed at such an angle that the test objects were in position to the right and in front of the observer.

A black tube was introduced between the eye of the observer and the small mirrors on the star wheel. A black cardboard screen was mounted on this tube so that practically no light except that reflected by the mirror at the end of the tube reached the observer's eye. The tube served two important functions: (1) The eye of each successive observer had to be placed in exactly the same position in order to be able to see the test objects; thus the test objects did not have to be realigned for each observer, and the angles of reflection were constant throughout the investigation. (2) The tube also facilitated the placing of the test objects at the beginning of each test period. An adjustable chin rest, fastened to the edge of the table, was provided to assist the observer in keeping his head immobile during the experiment.

The selection of a suitable test object was to some extent dependent upon the nature of the apparatus and also upon the procedure to be followed. The function of the first object exposed to the observer was merely to provide a fixation point but, in order to insure precise accommodation to this first object, it was necessary that this be a complicated geometrical form (such as a letter of the alphabet) which would require more precise accommodation than just a point. Consequently, block capital letters, subtending a visual angle of 5′ (the Snellen standard of visual acuity), were selected for this purpose.

The second test object was the crucial one of the pair. This object had to be a geometrical form which could be named by the observer or the position of which could be designated (as with the Landolt ring). Block capital letters of the same specifications as those used for the first test object were selected.

The distances of these two test objects from the eye of the observer were arbitrarily set at 4 m for the far object and 1 m for the near object. The plan of the investigation included the determination of the velocity of both near-to-far and far-to-near accommodation. Therefore, one set of letters of the appropriate specifications at 4 m served as the crucial test objects in near-to-far accommodation and as suitable fixation points for far to near accommodation. Likewise, the letters at 1 m distance were used as fixation points in near to far accommodation and as the crucial objects in far to near accommodation.

The tangent of the angle 5′ at 1 m is 1.45 mm; this is the size, then, of the letters to be exposed at a distance of 1 m. Similarly, the tangent of 5′ at 4 m is 5.80. Inland Gothic type in capitals, which is 1.45 mm high in 6 point, and 5.80 mm high in 24 point fulfills these requirements. The selection of the letter series to be used follows the results of an investigation by H. Banister (2), who found that there were two series of letters, five in each series, which were approximately equally confusing within the series. One of these is: O, C, D, Q, G; the other: H, K, M, W, N.

These ten letters, in 6 point size, were mounted, all with the same amount of surrounding white space (1 sq. cm) on a heavy cardboard disc 10 cm in diameter. In front of this disc was a sheet of dark grey cardboard containing an opening 1 cm square which was placed in such a position that one letter at a time could be exposed at the aperture by turning the disc on its pivot. The larger letters were mounted in a similar fashion, the disc, aperture, and cover being approximately four times as large. Paper and cardboard used in both these mountings were from the same pieces of stock and thus were of uniform quality. The illumination on these test objects was provided by light bulbs placed in inverted globes 5 inches in front of the letters and 7 inches diagonally below them. No other illumination was present in the room.

Procedure.—In a preliminary series of experiments, both of Bannister's sets of letters were used, but each was used as a separate series. Since the differences in accommodation time for the 'round' letters and the 'square' letters were not of interest in this investigation, only the series of 5 'square' letters was used for the subjects whose results are given below.

Before beginning on the actual determination of the accommodation time, it was necessary to ascertain whether the subject was able to read all of the letters correctly under the conditions of this experiment. The subject was requested to pronounce the letters appearing in the tube as they were changed by pivoting the discs containing the letters. If the vision of the subject was satisfactory, the determination of the accommodation time followed.

The subject was instructed somewhat as follows:

> "As the apparatus is put into operation, you will see two letters exposed. The first letter will be exposed for quite a long time and the second letter will be exposed for a shorter time. When the first letter appears, look at it carefully and fixate on it as clearly as you can. Continue to fixate on it until it is removed and the second letter is exposed. Look at the second letter and attempt to read it before it is taken away. After the second letter has disappeared, please pronounce it. It is not necessary to pronounce the first letter, just the second one. When the first letter appears, adjust your head so that the letter appears exactly in the center of the tube where it can be seen most clearly. Hold your head in this position until the second letter has appeared and disappeared. The second letter will appear in exactly the same position as the first."

Both the fixation letter and the first critical letter were set for exposure; the apparatus was set for a relatively long critical exposure and put in motion. The critical letter was changed after each exposure and the fixation letter was changed after each series of 5 exposures. The length of the critical exposure was reduced for each series. When two readings, differing by 5 degrees, were obtained, one of which showed more than 50 percent wrong responses and the other showed more than 50 percent right responses, an additional 5 observations were taken at each speed. If these additional 5 observations at each speed were respectively above and below 50 percent, it was assumed that the threshold lay between these two speeds.

This procedure was repeated with the position of the subject's head shifted so that the other eye could be used. The procedure was exactly the same as that for the eye used in the first threshold determination. In neither case was the eye which was not in use kept shut or occluded, since the black cardboard screen kept practically all distracting light from the eyes and produced the same effect as a blindfold on the unused eye. The procedure described above completed the first sitting. It is believed that in no case was either eye used long enough to produce fatigue.

At the second sitting, the first day's results were available to the operator and in most cases served as a guide in the location of the accommodation time threshold for the opposite direction. Otherwise, the procedure was exactly the same as for the first sitting.

At the completion of each sitting or series of sittings, the speed of revolution of the apparatus was calibrated. The unit of measure used in the setting of the two contact arms was 5 degrees. The temporal equivalent of the 5 degree unit span was determined from the daily calibration. This temporal unit varied between .060 sec.

and .065 sec. It may be noted here that the speed of revolution did not vary appreciably during the course of any one set of observations.

When the times for the readings above and below the threshold had been calculated, the 50 percent point were interpolated and this value was recorded, in thousandths of a second, as the threshold value.

Finally, this measure was corrected by subtracting one unit of time to allow for movement of the critical exposure mirror into position. It is believed that this final threshold value is a measure of the speed of 'perceptual accommodation,' which, in addition to the actual change in shape of the crystalline lens, includes the retinal lag, and the conduction time elapsing between the instigation of the nervous impulse in the retinal apparatus and the innervation of the ciliary muscle. The verbal response which follows, unlike the finger response in the reaction technique, is not included in the timing of the accommodation movement, although the fact that it is to follow the second exposure probably has some effect on the whole process.

Results

The results obtained from the 14 subjects whose accommodation time was measured are given in Table I. Four thresholds are given for each individual: The left eye and right eye for both near to far and far to near accommodation. The means of the columns are given below the raw data. The means of both eyes combined for near to far and far to near accommodation also are given.

The means of the seven highest and seven lowest cases in each column also are given. These indicate that the distributions of these four measures of accommodation time are similar. The means of the slow groups are approximately twice as great as those of the fast groups.

The fastest time recorded was that of Subject 3 with the right eye in far to near accommodation. This was .068 sec. This is about ten times as fast as his left eye in the same direction—.620 sec. The slowest time obtained was for the left eye of Subject 9 in near to far accommodation. This time was .691 sec.

An examination of the individual records on all four measures shows that sometimes one eye is much faster than the other (Subject 3) and also accommodation in one direction may be much faster than in the other direction (Subject 8). Comparisons of the distributions, however, indicate that individual differences are more pronounced than those among the four measures for the individual subjects.

TABLE I

ACCOMMODATION TIMES OF 14 SUBJECTS

N—near; F—far; L. E.—left eye; R. E.—right eye. Time recorded in seconds.

Subjects	N–F		F–N	
	L. E.	R. E.	L. E.	R. E.
1	.447	.447	.508	.403
2	.569	.564	.480	.272
3	.450	.285	.620	.068
4	.255	.270	.380	.280
5	.227	.508	.427	.275
6	.273	.236	.173	.205
7	.158	.173	.480	.544
8	.641	.641	.345	.260
9	.691	.233	.329	.442
10	.210	.270	.341	.279
11	.213	.275	.279	.452
12	.232	.279	.153	.397
13	.217	.145	.410	.347
14	.240	.176	.093	.163
Means	.345	.322	.358	.313
	.333		.336	
Means high seven cases	.475	.428	.472	.409
Means low seven cases	.214	.215	.245	.217

DISCUSSION

There were two conditions of method present in the previous investigations which it seemed desirable to avoid in the present research. One of these, and probably the most important with respect to the accuracy of measurement, was the necessity of a finger reaction by the observer to establish a criterion of accommodation and to time the movement.

There are several sources of error in the differential reaction technique. One of these is the necessity of the observer to establish a criterion of distinctness. To respond accurately to the same degree of distinctness at each exposure is an impossibility. Another is the necessity of a finger reaction to the visual stimulus. The speed of reaction to a relatively simple stimulus situation is known to vary considerably;

we may therefore assume that it would vary even more in a more complex situation such as that involved here. Thus, any attempt to correct for finger reaction time would be inaccurate. Furthermore, the use of the reaction method is limited to situations resulting in relatively long accommodation times, as Barrett was forced to admit. Where it is desired to measure accommodation of short distances, especially when the range being investigated is some distance from the eye and the dioptric change is correspondingly low, this method is entirely inadequate.

When all of these factors are considered in relation to the measurement of accommodation time obtained, even though an attempt is made to correct for the sources of error present, it is clear that the reliability and validity of the measure are doubtful. The alternative is to replace the differential reaction method by objectively controlling the time of exposure and establishing an objective criterion of accommodation. Ferree and Rand (**7**) and Banister and Pollock (**3**) did this by two different methods. Still another, described above, has been used in this experiment.

The other condition was that involving a change in direction of fixation as well as a change of accommodation. By the use of mirrors it was possible to eliminate this factor by presenting both test objects in exactly the same line of regard. This method is to be preferred to that involving shutters since the mirrors can be moved more quickly and accurately, without vibration, than the test objects themselves.

Several kinds of test objects have been used in past investigations of accommodation time. These include diagonally placed threads, pins, large and small 'O,' 'E' in four different positions, and other letters. The plan for this experiment was to use a series of letters which are easily confused. By this method, it is necessary to discriminate minor differences between the successive exposures; in order to be able to do this, quite accurate accommodation is necessary. When it was found that the round series of letters was more difficult for most individuals than was the square series, only the latter (H, K, M, N, W) was used.

After several subjects had participated in this experiment, it was evident that the test letters were not of equal difficulty. This had been anticipated on the basis of the frequencies of confusion given by Banister but appeared even more marked under these conditions. The W was the easiest for most individuals to read while the others, in order, were M, K, N, and H. This situation did not seriously affect the threshold determination, however.

The fact that the same letters were used for the fixation letters as were used for critical letters may have had some effect upon the responses. The motor aspect of perception, which is being demonstrated with increasing frequency, especially in vision, may influence the perception of the observer of the second exposure. Although there is no evidence in the records of these observers, due to the nature of the recording method, it is not unreasonable to suppose that the nature of the first letter exposed might have some effect upon the response in regard to the second letter.

The above considerations were suggested by the fact that a few of the observers developed 'fixations' on certain letters. One of these was H, another was W. No matter which letter was exposed, the observer almost invariably replied with the same letter. This condition appeared only when the length of exposure was near the accommodation threshold.

The fact that there were only 5 possible responses (and the subjects soon became aware that there were only 5 different letters) may have been a source of chance error. An attempt was made to overcome this, however, by exposing some of the letters twice in the same series without recording the response to the second.

It is impossible to state what kind of a test object is best for this purpose. Different letters, as used here, are probably the most convenient. The Landolt ring or a letter which can be placed in different positions presents difficulties in naming or designating the position. This question must be decided by experimental results.

The possibility of a chance error suggests that the threshold criterion should be higher than 50 percent correct re-

sponses. This percentage criterion is purely arbitrary; Ferree and Rand, for example, have used 3 out of 5, or 60 percent, correct responses as a criterion.

The 'tempo' of the apparatus is another undesirable factor. In a series of observations at a given speed, there is probably some adaptation to the tempo of the apparatus, so that the subject may establish a set which may even result in an accommodation movement before the actual appearance of the critical letter.[1] This would tend to reduce the observed accommodation time. This same condition was observed by Vierordt (**12**). He pointed out that an isolated accommodation movement took place more slowly than the same movement preceded and followed by similar movements in a series.

The second means of checking the apparatus is the examination of the results obtained from the subjects who participated in the investigation. Referring to Table I, it is clear that these data do not show any difference in the accommodation times of the two eyes nor is there any evidence that near to far accommodation is slower or faster than far to near. In regard to the first point, there is no reason to expect to find any difference between the left and right eyes of a group.

According to the previous investigations of far to near in relation to near to far accommodation, differences in time required for these movements were found. Barrett (**4**) was unable to measure near to far accommodation because it was too rapid; Seashore found that near to far was slower than far to near accommodation. Neither investigator attempts to establish a principle on the basis of his results; for one thing, the number of subjects in each case is entirely inadequate to serve as a basis for definite conclusions. The differences of the means (Table I) for the two eyes—.345 − .322, and .358 − .313—are much greater than the differences of the means for the two directions of accommodation—.358 − .345,

[1] To demonstrate that accommodation time was actually a factor in the determinations, several readings were taken with both test objects the same distance from the eye. This speed was then subtracted from the original reading. In 7 out of 8 cases, the resulting corrected time was larger than the correction and also larger than the unit of measurement—.060 sec.

.322 − .313, and .336 − .333. Since the number of subjects used here (14) is much larger than any previous group, these conclusions are probably more reliable than any presented previously. However, the distances and parts of the accommodation range used in the different investigations are not comparable. It is entirely possible that, whereas there are no differences in time for the two directions between 1 m and 4 m, there might be definite demonstrable differences at other distances and other parts of the range.

Barrett (4) presents an interesting conjecture in this connection. He reasons that near to far accommodation will be much more rapid when the near point is close to the punctum proximum of accommodation than when it is farther away because, when accommodation is maximum, the ciliary body and the lens capsule are at greater tension than when the eye is accommodated to a farther point, tends to increase the speed of this movement. Thus accommodation from 15 cm to 3 m may be faster than that from 1 m to 3 m. This same principle might apply to myophia and hypermetropia. If one individual has relaxed accommodation at 4 m (125 D myopia) while another is exerting .25 D or more (hypermetropia) at the same distance, this situation may make some difference in the accommodation time to another distance.

Except for the implications of the above considerations, it is doubtful whether it is reasonable to expect to find any consistent difference between near to far and far to near accommodation in the dynamic interrelations of the ciliary body, zonule, lens, and lens capsule. The interdependence of this complication of factors is difficult to evaluate; it seems reasonable that, in different individuals, the elastic portion of the system may be either stronger or weaker than the contractile mechanism, or both may be strong or weak. These conditions might be distributed at random, which would account for the discrepancies among previous investigators using small samples as well as the absence of difference in the results presented here. The 'young adult' was stipulated since the changes in presbyopia are usually of a definite character

in all individuals, the most striking characteristic being the recession of the punctum proximum.

In connection with this discussion, it might be well to refer to an observation by Fincham (**8**) in regard to the elastic portion of the system. The lens capsule, when treated with a specific elastin stain (such as orcein) becomes deeply stained while the lens itself, under similar treatment, is scarcely affected. This indicates that it is principally the lens capsule and not the lens which provides the elasticity. This is borne out by the observation that, when the capsule of an extirpated lens is cut, the opening spreads and the lens substance protrudes from the opening. This indicates that the body of the lens distends the lens capsule. Furthermore, the capsule is thicker at the periphery than in the center and thinner on the anterior surface than on the posterior. As a result, the lens, at rest, is thicker in the middle than at the edges and bulges more on the anterior surface than on the posterior.

The development and perfection of a method by which accommodation time can be measured accurately and quickly is of importance as such, but it may also be applied in other connections. A systematic study of accommodation time may, for example, lead to a better understanding of the accommodation process. Also, this method could profitably be applied to the study of presbyopia. If persons with presbyopia could be obtained for the purpose of studying accommodation time, and especially to observe the relation between the elastic system and contractile system, we might thus gain a clearer conception of what actually takes place in presbyopia.

The speed of accommodation is also of practical importance. Banister and Pollock (**3**) cite the application to athletics which involve 'keeping your eye on the ball.' Two of their subjects, differing in athletic ability, also differ in accommodation time, the athletic having the lower time. This might be worthy of systematic study.

This method is also applicable to studies of fatigue in alternate near and far accommodation. It is thought that this process, for example, is one of the most important factors in fatigue in reading, where a change in fixation from the read-

ing material to objects in the environment usually necessitates a large change in accommodation. This applies to other types of close work as well.

An examination of the four thresholds obtained from each individual might give some indication of the condition of the accommodation mechanism in each eye and also of the nature of minor refractive errors.

Conclusions

In the technique used the action of the extra-ocular muscles (of the eye being used) apparently has been eliminated since changes of fixation and convergence do not accompany the accommodation movement. Also, an objective criterion of accommodation has been incorporated into the device. This criterion, although apparently reasonably successful, can be improved upon.

The results of the subjects used in the investigation indicate that, in a group of persons, the average accommodation time for the left eye is approximately equal to that for the right eye.

For the distance at which the test objects were placed in this experiment (1 m and 4 m), there is no difference between the time required for near to far and for far to near accommodation.

Inspection of the individual readings of right and left eyes indicates that the fact that there is no reliable difference in means for left eye and right eye is a function of probability. For far to near accommodation, 13 subjects (about 93 percent) show a difference in time for right and left eye of more than .060 sec (the approximate unit of time measurement). For near to far accommodation, this same situation is observed in 50 percent of the cases.

(Manuscript received February 20, 1939)

Bibliography

1. Aeby, Die Accommodations-Geschwindigkeit d. menschlichen Auges, *Z. f. rat. Med.*, Dritte R., 1861, 9.
2. Banister, H., Block capital letters as tests of visual acuity. *Brit. J. Ophthal.*, 1927, 7, 49–62.

3. Banister, H. and Pollock, Accommodation time, *Brit. J. Psychol.*, 1929, **19**, 394.
4. Barrett, J. W., The velocity of accommodation, *J. Physiol.*, 1885, **6**, 46.
5. Copeland, R. S. and Ibershoff, A. E., *Refraction*, Phila.: Boericke and Tafel, 1906.
6. Duke-Elder, W. S., *Text-Book of Ophthalmology*, London, H. Kimpton, 1934, V. I, 742–752.
7. Ferree, C. E. and Rand, C., Speed of adjustment, *Amer. J. Psychol.*, 1919, 30, 40–61.
8. Fincham, E. F., The changes in the form of the crystalline lens in accommodation, *Amer. J. Physiol. Optics*, 1926, **7**, 469.
9. Johanson, A. M., The influence of incentive and punishment upon reaction time, *Arch. Psychol., N. Y.*, 1922, No. 54, p. 35 and pp. 36–40.
10. Seashore, C. E., Monocular accommodation time, *Yale Univ. Stud.*, 1893, **5**.
11. Tefft, L. E. and Stark, E. K., Speed of accommodation, *Amer. J. Ophthal.*, 1922, **5**, 339–342.
12. Vierordt, Versuche über die Zeitverhältnisse des Accommodationsvorganges im Auge, *Arch. physiol. Heilkunde*, N. F., **1**.
13. Volkmann, W. F., *Wagner's Handwörterbuch d. Physiol. 'Sehen,'* B. III, S. 309, 1846, Braunschweig.
14. Wilmer, W. H. and Berens, C. *Eye in Aviation.* Aviation Medicine in A. E. F. Gov't Print. Off., Washington, D. C., 1920.

THE FIXATIONAL PAUSE OF THE EYES

BY DONALD C. ARNOLD AND MILES A. TINKER
University of Minnesota

Two important aspects of oculomotor behavior in reading are fixational pauses and the interfixation or saccadic movements. It is well known that reading perception occurs only during the pauses since there is no clear vision during the saccadic moves. Pauses are relatively long in duration in comparison with movement time. In two studies Tinker (**9**, **12**) has shown that the percentage of the reading time taken by the pauses varies somewhat with kind of material read. With very easy prose the pause time is as low as 90.5 percent, but with very difficult material such as mathematical formulas it is as high as 97.2 percent. Apparently in the more complex types of reading, in which the mental processes of apprehension and assimulation play an important part, relatively more time is devoted to the pauses.

To what extent is pause duration extended beyond that required for clear perception of the symbols involved? Erdmann and Dodge (**3**) found that perception was possible with exposures as short as 0.25 ms. (millisecond). Various other investigators have employed brief exposures ranging from 0.25 to 70 ms. Dodge (**2**) has pointed out that perception with these short exposures is far divorced from perception in the normal reading situation. In fact, to get perception at these short exposures it was necessary to use dark pre-exposure and post-exposure fields to enhance the positive after images. In everyday situations, such as reading, stimulation during normal fixation pauses is never provided with dark pre- and post-exposure fields. Dodge (**2**) worked out an exposure technique for successive presentation of words in a series that was analogous to the normal reading situation. Although Dodge did not recognize the fact, his situation was not strictly comparable to normal reading since eye movements

need not occur in his set-up. He found that 80 to 100 ms. is the minimum adequate exposure for perception. Although an exposure of 60 ms. permitted the correct reading of single words, an exposure of 100 ms. was necessary for a fairly cleared-up impression. In the clearing-up process the effects of past stimulation fades out and the new sensory impression reaches its full maturity so that clear perception is possible. On the basis of these studies Dodge comes to the conclusion that "Both experimentally and practically the operation of the visual clearing-up process precludes a succession of adequate visual fixations under .1″ each" (2). In the continuous reading situation we have the problem of how rapidly adequate fixation pauses may succeed one another. If a pause has a duration of between 80 and 100 ms. it will presumably furnish minimum adequate exposure for perception. When the pause is between 60 and 80 ms., it would probably be inadequate except in special circumstances such as complete familiarity with the text or when a symbol has been inadequately perceived by one fixation and needs the succeeding brief fixation to furnish supplementary information. When the fixation pause has a duration of more than 100 ms., however, the time in excess of the 100 ms. is presumably due to other factors than the requirements of clear perception. Such factors might include reaction time of the eye, requirements of apprehension and assimulation, the minimum time required for the eye to pause and then move on to the next fixation, and the speed with which binocular coordination is reestablished after a saccadic interfixation movement.

The experiment conducted by Cobb and Moss (1) attempted to establish the minimum duration for the fixational pause of the eyes. Their procedure consisted essentially in having the subject move his eye from one fixation to another 20° away with a stop half way across. They reported the number of light spots seen at the central fixation. Since these spots appeared in succession at known intervals the duration of the fixation pause was indicated in terms of the number of spots reported. The mean pause duration values accepted as adequate by the authors ranged from 113 to

167 ms., with the median subject at 140 ms. These data cannot be accepted as representative for two reasons: (1) The report required is highly subjective. To perceive accurately and report correctly are both subject to error. For instance, an over or under estimation of one spot of light means an error of 12 to 28 ms. according to the timing of the spots. The direction of the error, however, would probably be toward underestimation. (2) Monocular vision was employed. The results cannot be applied to performance during reading where binocular vision is employed. Stromberg (**8**) has shown that convergence occurs during interfixation movements and divergence during the pause. It is possible that the readjustment of divergence during the pause complicates the clearing-up process.

In the present experiment the purpose is to determine, by the photographic method, the minimum duration of the fixational pause during binocular vision. Situations investigated include the time needed (1) to fixate long enough to perceive a letter, (2) to fixate a dot, and (3) to fixate anywhere in a blank area between the two end fixation points.

The Minnesota eye-movement camera (11) was used. The reflected light was interrupted 60 times per second. Each dot and each space between dots on the photographic record, therefore, represented 1/120 second. Possible error in reading any pause duration was approximately 4 ms.

In situation *A*, three dots four inches apart were lined up horizontally on a white cardboard. The four-inch distance between dots represented approximately 12 degrees of visual arc. The subject was instructed to fixate on the dot at the left, then at a given signal to move the eyes across the card to the dot at the center and stop there, and then to move on as quickly as possible to the fixation dot at the right. Fixation was maintained at the right until the subject was told to return to the left hand dot.

In situation *B*, the two end dots were the same as in *A* but at the central location there was a lower case letter. The subject was instructed to stop there just long enough to identify the letter and then to move quickly to the right dot and then to report the name of the letter. The uncovering of the stimulus card was the signal to move from the left to the center.

In situation *C*, there were two dots 8 inches apart horizontally. The subject was instructed to move from the left to any position in the general region of the center, to stop there and then to move on as quickly as possible to the right dot. Presumably these situations required the subject (1) to make the briefest possible pause in a definite location; (2) to pause long enough to perceive a single symbol, and (3) to make the briefest possible pause without fixating upon a definite object.

Ten university students were employed as subjects. Each observed for two sittings. At each sitting the subject was given a few practice trials and then a series

of approximately 36 trials in each situation (*A*, *B* and *C*). There was a brief rest of a few seconds between trials and a longer rest at the end of each series. In condition *B*, a different card with a different letter was used on successive trials. To equalize practice effects the *A B C* sequence of situations was used at the first sitting and the *C B A* at the second.

Results and Discussion

The basic data of the study are given in Table I. Individual as well as group results are shown. As seen in the

TABLE I

Measurement of Pause Duration

The mean, N, and SD are indicated for each situation: A, fixation of a dot; B, fixation of a letter; and C, fixation within an area. Individual and group data are presented. Durations in ms.

Situation	Measure	Subjects										Total Group
		A	B	C	D	E	F	G	H	I	J	
A: Dot	N	69	68	73	71	62	75	73	70	74	36	671
	M	188	208	197	189	141	125	116	183	276	125	172
	SD	28	60	54	62	21	32	26	41	43	44	40
B: Letter	N	78	53	77	73	44	75	71	67	64	65	667
	M	147	115	147	175	134	146	135	110	239	217	157
	SD	24	43	25	32	32	42	25	51	54	49	41
C: No Object	N	69	39	68	68	63	72	60	71	71	66	647
	M	152	218	263	213	140	185	152	203	275	173	199
	SD	37	54	61	32	33	42	30	60	53	42	52

N-row, there was some variation in number of pauses from subject to subject. This was due to illegible records in some instances. Note that in fixating the dot the mean minimum pause durations for individuals range from 116 to 276 ms. The average for the group is 172 ms. It takes less time to fixate long enough to perceive a letter (situation B) than to briefly fixate a dot. Individual means range from 110 to 239 ms. with a group average of 157 ms. This tendency to take less time in situation B than A is consistent for 8 of the 10 subjects. In situation C where there was no object to fixate the minimum pause duration is greater. The group average is 199 ms. In 7 out of 10 cases the individuals took more time in C than in A while for two subjects the means are approximately the same. When situation C is compared with

B, 9 of the 10 subjects have longer pause durations in C. The differences between the group means are highly stable since D divided by SE_D is 7.2 for A vs. B, 14.1 for B vs. C, and 9.7 for A vs. C.

Readjustment fixations were fairly prevalent in all situations. In fixating the dot (A), 64 percent of the pauses involved readjustments; in fixating the letter (B) 75 percent; and where no object was present (C), 55 percent. To what extent is the pause duration affected by these readjustments? The answer is found in Table II. On the average, fixations

TABLE II

PAUSE DURATION AS INFLUENCED BY READJUSTMENT

Group results for ten observers are given. The number of pauses and mean pause duration for each condition are listed (Durations in ms)

Situation	Without Readjustment		With Readjustment		Both	
	N	Mean	N	Mean	N	Mean
A: Dot	431	161	240	178	671	172
B: Letter	500	153	167	164	667	157
C: No Fixation	360	181	287	208	647	199

without readjustment are somewhat shorter in duration than those with readjustment. Although these differences are not large they are probably stable, especially in situations A and C.

In Table III is given the part played by the initial and final parts of the fixational pauses involving readjustment. The mean duration of the initial part of the pauses varies from 68 to 88 ms. All of these durations fall below Dodge's

TABLE III

MEAN DURATION OF INITIAL AND FINAL PART OF READJUSTED PAUSES

(Durations in ms.)

Situation	N	Initial Part	Final Part*	Total Pause
A: Dot	431	69	109	178
B: Letter	500	68	96	164
C: No Fixation	360	88	120	208

* In a very few instances there were two readjustments instead of one. In such cases the last two parts were counted together since no appreciable error was involved.

specifications (100 ms.) of exposures that yield fairly cleared-up visual impressions. In only situation C is the mean duration of the initial part of the pause long enough for adequate perception, *i.e.*, it falls between 80 and 100 ms. The final part of the pause durations ranges from 96 to 120 ms. Note that these latter durations are either just above or just below Dodge's 100 ms. minimum limit noted above. Can perception in situations A and B take place only during the final part or does it begin in the initial part and complete itself in the final? There are indications that the beginning of the perception occurs in the initial part and completes itself in the final part. In the first place the duration of the final part is shorter than is characteristic of unadjusted pause durations, which are 161, 153 and 181 ms. respectively for A, B and C (Table II). Furthermore, the combined duration of initial plus final parts of readjusted pauses is only slightly longer than that of unadjusted pauses. It would seem, therefore, that when readjustment of a fixation pause occurs, the initial and final phases are parts of a unitary whole. If an act of perception is required, it probably has a beginning in the initial part but is only completed during the final part of the fixation duration.

Discussion

The significant differences in minimal duration of pauses when fixating a dot, perceiving a letter, and when merely stopping as briefly as possible with no object to fixate are interesting. One might suggest that less exact fixation was required to perceive the letter than when fixating the dot since the former could be identified even though some distance from the center of the fixated area. Nevertheless there were more readjustments in situation B than in A. The relatively long durations in situation C where there was no object to fixate may be due partly to difficulties of orientation at the central area before moving on. Subjects commented that this was the most difficult of the three tasks. Fewer readjustments occurred in C, however, than in either A or B where the durations were shorter. Note that frequency of read-

justed pauses does not change the trend of pause duration in A, B and C. Table II shows that the trend is the same whether the pauses are readjusted or not. Since the eyes converge during saccadic movement and diverge during fixation (**8**), it may take longer to re-establish coordinate vision in a pause without a definite fixation object (C) and thus delay the clearing-up process and orientation to the next move.

The results of Cobb and Moss (**1**) reveal mean pause durations that range from 113 to 167 ms. with a median value of 140 ms. for 9 subjects. The results which the authors emphasize most yield a group mean of approximately 150 ms. Reference to Table I shows that the pause durations in our study are considerably longer, especially in situations A and C. This suggests that inaccuracies occurred in the Cobb and Moss values, probably due to the subjective technique employed to determine pause duration.

It is possible that the reaction time of the eye is a determinant of minimum pause duration. Simple eye reactions are in some ways similar to the response when the eye makes the briefest pause possible. Hackman (**5**) found the mean for simple reaction time of the eye to be 173 ms. This value is approximately the same as the mean fixational pause of 172 ms. on the dot.

Reference back to Table I reveals that the mean pause duration for every subject in all three fixation situations is longer than the 100 ms. specified by Dodge (**2**) as the minimum for a well cleared-up visual process. In general, therefore, the briefest pause the eye can make appears to have a duration which is more than sufficient for clear vision. Examination of the single responses of the subjects revealed only about 5 percent with durations less than 100 ms. Of these, the very shortest was 75 ms. but most of them fell between 80 and 100 ms. which Dodge specifies as the minimum adequate exposure for perception. None was less than the 60 ms. which Dodge found sufficient for correct reading of single words although clear vision was absent.

Minimum duration of fixational pauses as here measured

may be compared with pause duration in reading. In Table IV are given data from several experiments. In reading easy prose the *group* mean pause durations in several investigations range from 217 to 246 ms. When hard prose is read the range is 230 to 277 ms. In special types of reading such as algebra and objective questions the pause durations range from 272 to 404 ms. In reading Chinese they are about 300 ms. In a group of 66 subjects reading hard prose the *individual* mean durations ranged from 174 to 310 ms. (4). All of these group mean durations are considerably longer

TABLE IV

MEAN PAUSE DURATION IN READING

(Durations in ms.)

Author	N	Kind of Material Read	Pause Duration
Walker (15)	50	Easy prose	246
Tinker (14)	64	" "	224
Tinker (13)	77	" "	217
Sisson (7)	60	" "	227
Walker (14)	50	Hard prose	277
Tinker (10)	16	" "	254
Tinker (13)	77	" "	230
Sisson (7)	60	" "	236
Frandsen (4)	66	" "	243
Tinker (10)	16	Algebra problems	272
Tinker (10)	10	Chemistry text	276
Tinker (10)	15	Algebra formulas	404
Frandsen (4)	66	Objective questions	294
Shen (6)	13	Chinese: vertical	305
Shen (6)	13	Chinese: horizontal	294

than the mean pause durations discovered in this study. It is obvious, therefore, that fixational pauses in reading have mean durations that are longer than the briefest pause the eye can make and much longer than the minimum exposure needed for adequate seeing.

It is true that subjects make single pauses that very considerably from the means given in Table IV. The photographic records of two subjects who read the 9 kinds of material described by Tinker (12) were examined. Subject I made a total of 815 pauses, only 14 (1.7 percent) of which had durations less than 100 ms. The range was 82 to 618 ms. The second subject made 767 pauses, only 13 (1.7 percent) of

which had durations less than 100 ms. The range was 72 to 498 ms. These figures indicate that although readers make a few individual pauses that approximate the minimum exposure adequate for visual perception, most fixations have durations which exceed greatly that required for a well cleared-up visual process. When pauses as brief as 72 ms. are made it is probable that the material had been seen in peripheral vision during the previous fixation and the 72 ms. were sufficient to supplement and complete the perception that had already begun earlier.

Why then are the longer pause durations of 200 to 600 ms. frequently employed in reading? Reading consists of more than seeing clearly the printed symbols. The subject must comprehend and assimulate the ideas and relationships involved. The more complex and difficult these are, the longer the pause durations to permit adequate comprehension and assimulation before moving on to new material. This is well illustrated in reading algebraic problems, formulas, and objective questions where the pauses frequently endure for about one-half second.

Summary and Conclusions

1. The briefest possible pauses made in fixating a dot, a letter and stopping in a blank area were measured by the photographic method. Ten subjects were used.

2. On the average it took 172 ms. to fixate the dot, 157 ms. for the letter and 199 ms. to stop without an object to fixate. The trend of the individual means was highly consistent.

3. Readjustments during fixation were frequent in all three situations. Since the duration of the readjusted pauses was similar to that of the unadjusted pauses it is suggested that the two parts of the readjusted pauses supplement each other to form a coördinated unitary reaction.

4. The briefest mean pause is longer in duration than the exposure needed for a well cleared-up perception.

5. Factors that may be involved in determining the duration of the shortest pause possible include: (1) divergence of the eyes during the pause after the convergence that occurs in saccadic movements; and (2) reaction time of the eye.

6. Single fixations of the eyes seldom have durations less than 100 ms. and practically never less than the minimum exposure needed for adequate perception. This is true both for fixating a single object and for fixational pauses in reading.

7. Fixational pauses in reading have mean durations that are much longer than occur during briefest possible fixations as measured in this experiment. These longer durations in reading may be explained in terms of the requirements for comprehension and assimulation of the materials read.

(Manuscript received March 10, 1939)

References

1. Cobb, P. W., and Moss, F. K., The fixational pause of the eyes, *J. Exp. Psychol.*, 1926, **9**, 359–367.
2. Dodge, R., An experimental study of visual fixation, *Psychol. Monogr.*, 1907, **8**, No. 35, 1–92.
3. Erdmann, B., and Dodge, R., *Psychologische Untersuchungen über das Lesen auf experimenteller Grundlage*, Halle, 1898.
4. Frandsen, A., An eye-movement study of objective examination questions, *Genet. Psychol. Monogr.*, 1934, **16**, No. 2, 79–138.
5. Hackman, R. B., *An analysis of certain factors operating in the reaction of the eye to suddenly appearing peripheral stimuli.* Ph.D. Thesis in Psychology, University of Minnesota, 1938.
6. Shen, E., An analysis of eye movements in the reading of Chinese, *J. Exp. Psychol.*, 1927, **10**, 158–183.
7. Sisson, E. D., Habits of eye-movement in reading, *J. Educ. Psychol.*, 1937, **28**, 437–450.
8. Stromberg, E. L., Binocular movements of the eyes in reading. *J. Gen. Psychol*, 1938, **18**, 349–355.
9. Tinker, M. A., Eye movement duration, pause duration, and reading time, *Psychol. Rev.*, 1928, **35**, 385–397.
10. Tinker, M. A., A photographic study of eye movements in reading formulæ, *Genet. Psychol. Monogr.*, 1928, **3**, No. 2, 67–182.
11. Tinker, M. A., Apparatus for recording eye movements, *Amer. J. Psychol.*, 1931, **43**, 115–118.
12. Tinker, M. A., Time taken by eye movements in reading, *J. Genet. Psychol.*, 1936, **48**, 468–471.
13. Tinker, M. A., Reliability and validity of eye-movement measures of reading, *J. Exp. Psychol.*, 1936, **19**, 732–746.
14. Tinker, M. A., Motor efficiency of the eye as a factor in reading, *J. Educ. Psychol.*, 1938, **29**, 167–174.
15. Walker, R. Y., The eye movements of good readers, *Psychol. Monogr.*, 1933, **48**, No. 199, 95–117.

THE RELATION OF TIME ESTIMATION TO SATIATION

BY ARTHUR BERMAN

Institute of Child Welfare, University of California [1]

In recent studies [2] on satiation, time estimations were secured as a part of the main body of data. These time estimations seem to warrant reporting here, for they provide additional experimental evidence on the problem of the subjective perception of time and its relation to satiation. More specifically, an attempt is made to probe the oft repeated belief that time passes slowly when one is satiated, *i.e.*, a satiated individual always overestimates the temporal period required to become satiated.

The accumulated evidence concerning estimations of time by satiated subjects is equivocal. Hoche (**4**), Lossagk (**5**), and Wyatt *et al* (**8**) found that satiated individuals tended to overestimate time in such a state. Thompson (**7**), Barmack (**1**), and Lossagk (**5**) [3] found either no overestimation, or that time was actually underestimated. Swift and McGeoch's (**6**) subjects underestimated a 10 min period filled by copying nonsense syllables, dry sense material, and selections from *Huckleberry Finn.* Since the nature of the obtained estimations may be in part a function of the experimental design, we record here in detail the conditions under which our time estimations were obtained.

Method and Results

I

Satiation, as employed in this paper, is operationally defined. An individual is satiated when he rejects (at least more than once) an object or activity which was

[1] The experimental work reported in this paper was done at the University of Southern California. Special thanks are due Mr. Homer J. Hare for his assistance with Exp. II.

[2] Unpublished experiments by the writer.

[3] Under certain specific experimental conditions.

initially desired or pursued; and reports or exhibits coördinately characteristic 'attitudes' of satiation.

Twenty-two adults, between the ages 20 to 47 years, volunteered as subjects. The majority were students attending the summer session (1938) of the University of Southern California.

The subjects were segregated by chance into two groups—alternately as they appeared at the laboratory; seven females and four males composed the satiation group, six females and five males the non-satiation group. The functional difference between the two groups was that the satiation subjects continued to perform on the maze until two rejections (non-resumptions); the non-satiation subjects continued until the criterion of three successively correct trials on the maze was achieved.

Time estimations in these experiments were always secured verbally immediately following the completion of the task. The subject later recorded his judgement independently on a rating blank. This procedure served as a check on his oral statement. The subject was unaware that an estimation of his activity time was to be taken until the actual request was made. Because of this lack of awareness, and the somewhat lengthy duration of activity, it was impossible for the subject to employ a mechanical aid, *e.g.*, counting, etc.

A stylus maze was the medium for the induction of satiation. This maze has been thoroughly described elsewhere (2) so that only a brief description is given here.

> The maze is one of the punch-board variety and contains 361 choices arranged symmetrically in columns $\frac{3}{8}$ inch apart. Externally its appearance is of an inverted box with numerous holes in the surface; its size is 1 foot square. A small white lamp (6 volts) is set at the center head of the panel and signals correct responses. A metal plate is fixed below the surface panel, space being allowed for the insertion of a pattern. Holes were drilled in this pattern for a specific design, which, when inserted, corresponds with identical holes in the surface panel. Current was supplied by 4 dry cells; and 2 Cenco counters—one to measure total responses, the other correct responses—were introduced into the circuit. A standard laboratory stylus was used. A right triangle approximately 7 inches high was the design to be solved.
>
> The members of the satiation group were given the following directions:
>
> You see before you what is called a learning board. You are to take this stylus and punch these holes until you can duplicate a pattern which I will show you. [Demonstrate, and give practice using stylus.] If the light glows when you insert the stylus, that is a correct response, and that choice is in the pattern—you must get this again if you can. If the light does not light, that is incorrect, and you must not enter this choice again if you can help it. You are to continue until you can do the pattern without error, *and then keep on just as long as you possibly can*—for upon its completion a second much more difficult pattern will be given you to solve. I reiterate: continue just as long as you possibly can. When you feel that you cannot go on, you may stop. Remember, however, you will be penalized. Do your best as you are in competition with the rest of the group. Do you have any questions?[4]

---

[4] All of the subjects were subjected to the second pattern after time estimations were obtained. This procedure served as an incentive, and to maintain a high level of aspiration. The data of the second learning task is not of immediate concern here.

The subject was then given the rating blank reproduced below but was not permitted to see Part B.

Age———— Name————————————
Sex———— Number————————————

A

Encircle the word that *best* describes your feelings or condition. Be as accurate as you can.

(1) Extremely bored..bored..indifferent..eager..extremely eager
(2) extremely tired..tired..neither tired-energetic..energetic..extremely energetic
(3) muscles extremely tense..tense..not tense..relaxed..extremely relaxed

B

Encircle the word that *best* describes your condition the *moment* you finished working.

(1) Extremely bored..bored..indifferent..eager..extremely eager
(2) extremely tired..tired..neither tired-energetic..energetic..extremely energetic
(3) muscles extremely tense..tense..not tense..relaxed..extremely relaxed
(4) very attentive..attentive..slightly inattentive..distracted..very distracted

How long have you been working at this task?

Did you feel *mentally tired* at the moment you finished? *Explain.*

After the subject returned the rating blank—in the interim the pattern had been inserted into the maze—he was shown a copy of the pattern (right triangle) for 10 seconds, and then told to proceed. The subject continued until he stopped of his own accord, or signified that he wanted to stop; and then was told to keep on. When he signified for the second time his intention of discontinuing, or as in most instances refused to go on, the period was terminated. After the time estimation was taken orally, Part B of the rating blank was given.

There was no necessity of repeating methodological details for the non-satiation group. They differed only in the respect of the varied criterion of solution previously mentioned.

Prior to the task, 1 of the satiation subjects was 'extremely eager,' 7 reported 'eager,' 2 'indifferent,' and 1 'bored' (Table I). After completing the task, 2 reported 'eager,' 2 'indifferent,' and 7 'bored' (Table II). One of the latter two reporting 'eager' described the activity as monotonous in his introspection. This rating questionnaire as an index of satiation left

TABLE I

SUBJECTIVE ATTITUDE PRIOR TO TASK (EXP. I)

	Ext. Bored	Bored	Indiff.	Eager	Ext. Eager
Satiation.....		1	2	7	1
Non-Satiation.....		1	3	6	1

TABLE II

SUBJECTIVE ATTITUDE AFTER COMPLETION OF TASK (EXP. I)

	Ext. Bored	Bored	Indiff.	Eager	Ext. Eager
Satiation.....		7	2	2*	
Non-Satiation.....			4	7	

* One subject reported satiation in written introspection, although he had encircled 'eager.'

something to be desired, for once the subject completes his task, and must recall his satiation, he tends to minimize his former condition. The directions were, for certain individuals, equivocal—in spite of precautions taken—for those who reported 'eager' later said they were 'eager' for the following task and not for the one just completed.

The non-satiated individuals, in this same situation, reported, before the task: 1 'bored,' 3 'indifferent,' 6 'eager,' and 1 'extremely eager' (Table I). Following the task, 4 were 'indifferent,' and 7 'eager' (Table II). The task seems to have increased the desire in these subjects for further activity (at least in collective terms), whereas the satiation group reported greater satiation.

Before attempting the task 1 member of the satiation group said he was 'tired,' 6 were 'neither tired-energetic,' and 4 were 'energetic' (Table III). After completion of the

TABLE III

SUBJECTIVE FEELING OF FATIGUE PRIOR TO TASK (EXP. I)

	Ext. Tired	Tired	Neither Tired-Energ.	Energ.	Ext. Energ.
Satiation.....		1	6	4	
Non-Satiation.....		3	5	3	

task 1 was 'extremely tired,' 4 were 'tired,' 5 'neither tired-energetic,' and 1 'energetic' (Table IV).

In a similar situation, before the task, 3 members of the non-satiation groups were 'tired,' 5 'neither tired-energetic,' and 3 'energetic' (Table III). After the task, 1 was 'extremely tired,' 5 were 'tired,' 3 'neither tired-energetic,' and 2 'extremely energetic' (Table IV).

Both groups show increased fatigue with activity; but no greater fatigue is indicated by the satiated subjects than the non-satiates, although individual variations occur.

The evidence from the rating blank may safely be evaluated as indicating that the members of the satiation group reported themselves to be satiated in varying degrees; whereas the non-satiates report a lack of satiation. There was little difference in report of fatigue. Since individual variation in the groups was large, and we are dealing here with differences in degree, it is possible that some members of the non-satiation group may have approached a slightly satiated state. However, the differences in activity time between the two groups, and the qualitative reactions exhibited during performance, indicated a discrete difference in satiation between the two groups.

TABLE IV

SUBJECTIVE FEELING OF FATIGUE AFTER COMPLETION OF TASK (EXP. I)

	Ext. Tired	Tired	Neither Tired-Energ.	Energ.	Ext. Energ.
Satiation	1	4	5	1	
Non-Satiation	1	5	3		2

Table V summarizes the data from Exp. I. It is desirable in time estimation experiments to consider each individual estimation in its own right and to minimize statistical manipulation.

Eight of 9 (89 percent) satiated subjects (the estimations of two subjects were discarded because of unlooked for orientation by the university bell) underestimated the time required to become satiated. Seven of the 11 (64 percent) non-satiated subjects underestimated the time required for them to solve the problem (= 3 continuous correct trials). The exact time of activity, the estimated time, and the difference between the two, are given for each subject in Table V. In quantitative terms the mean activity time for the satiation group was 41 min.; the mean estimated time 31 min. 7 sec. The average time of activity for the non-satiates was 4 min.

52 sec.; the estimated time 5 min. 16 sec. An average underestimation of 12 min. 11 sec. for the satiated group was seen; a 30 sec. overestimation for the non-satiates.

In every instance, but one, the satiates underestimated the time required for them to become satiated. The group of non-satiates varied between over- and underestimation: 4 underestimating and 7 overestimating. The latter as a whole slightly overestimated their time of performance. The error

TABLE V

TOTAL TIME, AND ESTIMATED TIME, REQUIRED TO BECOME SATIATED (EXP. I) *

Satiation Group					Non-Satiation Group				
Subj.	Perf.† Time	Choices	Est.‡ Time	Diff.‖ P.T.−E.T.	Subj.	Perf. Time	Choices	Est. Time	Diff.‖ P.T. −E.T.
E.H.....	35′38″	1034	35′	− 38″	W.S.	6′33″	197	15′	+8′27″
Z.D.....	53′02″	2254	35′	−18′02″	B.C.	3′15″	131	3′	− 15″
M.H.....	16′21″	969	—§	—§	W.	5′52″	332	5′	− 52″
N.J.M...	78′11″	3395	38′	−40′11″	T.Mc.	4′52″	371	3′	−1′52″
J.Mc....	23′59″	1312	45′	+21′01″	W.G.	2′48″	164	2′	− 48″
M.E.T...	45′	1401	—§	—§	H.B.	3′25″	222	4′	+ 35″
H.G.....	57′01″	3739	30′	−27′01″	E.F.	3′51″	212	3′	− 51″
N.W.....	6′34″	408	5′	− 1′34″	H.H.	3′17″	132	5′	+1′43″
S.....	32′03″	1231	15′	−17′03″	L.M.	5′39″	196	3′	−2′39″
C.R.....	51′05″	4544	47′	− 4′05″	J.C.	5′55″	273	5′	− 55″
R.B.....	52′10″	4547	30′	−22′10″	B.B.	7′05″	347	10′	+2′55″
Mean....	43′18″	2496	31′07″	−12′11″	Mean	4′52″	234.2	5′16″	+ 30″

* Satiation group continued until satiated; non-satiation group to three continuous successful trials.

† Performance time.

‡ Estimated time.

§ Judgement discarded: objectively oriented by University bell signalling the hour

‖ Minus quantity denotes underestimation; plus quantity: overestimation.

of estimation (departure from objective duration) was much less for the non-satiates, since only a slight over- or underestimation was made by each member of this group (mean error = 30 seconds). The satiation group showed comparatively large errors of estimation (mean error = 12 min. 11 sec.). It is easily seen, in this connection, that the intervals to be judged were in nearly all cases greater for the satiation group, thus permitting wider variation (or error).

II

The experimental design of the second experiment follows closely that of the first, and was designed to test—under specifically varied conditions—the results found in Exp I. This precludes the necessity of repeating methodological details for Exp. II where they approximate those of Exp. I.

A ten alley, high-relief, U type finger-maze of the pattern L R R L L R R L L R Fig. 1) was employed in place of the punch-board maze previously described. Both

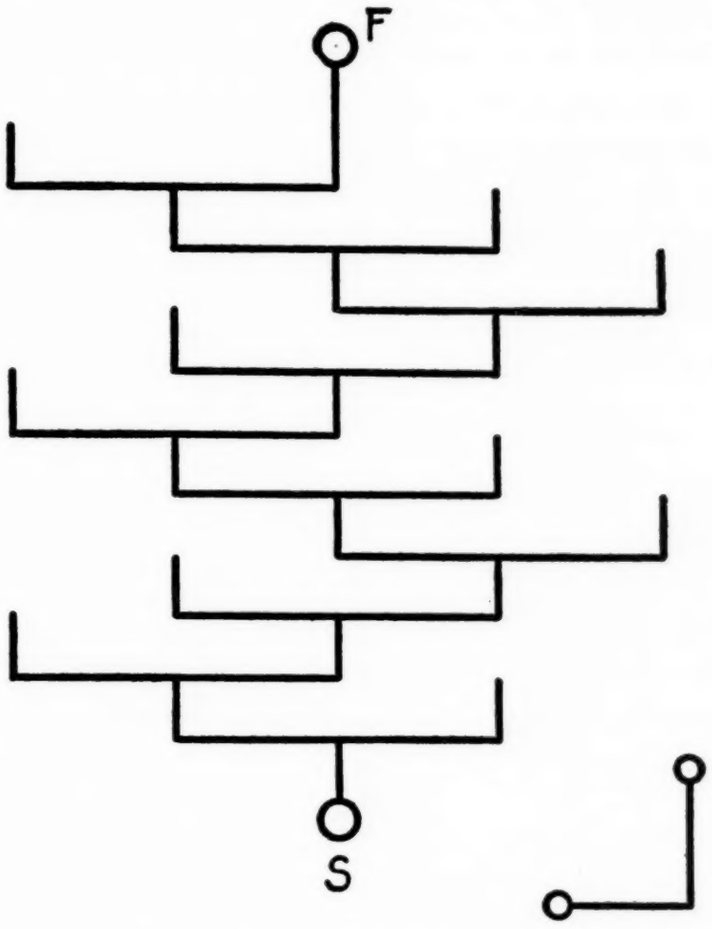

FIG. 1. The multiple U finger-maze.

satiation and non-satiation subjects performed identical tasks except that the former now continued to 3 non-resumptions of activity; the latter to 5 consecutively successful trials. An attempt was made to increase the duration of performance over that of Exp. I by extending the criterion of satiation and learning respectively.

Twenty-four adults[5]—13 females and 11 males—volunteered for the experiment. They were drawn from the same sample of population as described for Exp. I. In the satiation group there were 7 females and 5 males; in the non-satiation group 6 females and 6 males.

[5] Three of these subjects also participated in Exp. I.

The subject, on his appearance at the laboratory, was given Part A of the questionnaire previously described. He was then thoroughly blindfolded and told to 'feel' a small sample of the maze. This sample was not a part of the true maze. The subjects of the satiation group were then given the following directions.

"You are going to be given a learning problem called a maze. You will be blindfolded, and asked to trace a pattern with the index finger of your preferred hand. You are to continue working at the pattern just as long as you possibly can—even though you can do it without an error. A second pattern will be given you to solve sometime later. The longer you persist in the first problem, the easier will be the second problem. When you feel that you absolutely cannot continue the pattern that I am about to give you, you are at liberty to stop. Remember, however, that you will be penalized."

'When I give you the word, you may begin.'

Upon the completion of the activity, time estimation was taken, and the remainder of the rating blank administered.

Before the task was begun, 9 of the satiation group were 'eager' and 3 'indifferent' (Table VI). After the task, 3 were 'extremely bored,' 5 'bored,' and 4 'indifferent' (Table VII). For the non-satiation group, before the task, 1 was 'extremely eager,' 7 'eager,' and 4 'indifferent' (Table VI). After the task, 6 were 'eager,' 5 'indifferent,' and 1 'bored' (Table VII).

TABLE VI

SUBJECTIVE ATTITUDE PRIOR TO TASK (EXP. II)

	Ext. Bored	Bored	Indiff.	Eager	Ext. Eager
Satiation.....			3	9	
Non-Satiation.....			4	7	1

TABLE VII

SUBJECTIVE ATTITUDE AFTER COMPLETION OF TASK (EXP. II)

	Ex . Bored	Bored	Indiff.	Eager	Ext. Eager
Satiation.....	3	5	4		
Non-Satiation.....		1	5	6	

Similarly, before the task was begun, 2 of the satiation group were 'tired,' 5 'neither tired-energetic,' and 5 'energentic' (Table VIII). After the task, 2 were 'extremely tired,' 5 'tired,' and 5 'neither tired-energetic' (Table IX). Four of the non-satiates were 'tired' previous to the task, 6 'neither tired-energetic,' and 1 'energetic' (Table VIII).

Following the task, 2 were 'tired,' 7 'neither tired-energetic,' and 3 'energetic' (Table IX).

The rating blank—as in Exp. I—indicates that the shift in the report of the satiation subjects is toward the satiation end of the scale, whereas the report of the non-satiation group

TABLE VIII

SUBJECTIVE FEELING OF FATIGUE PRIOR TO TASK (EXP. II)

	Ext. Tired	Tired	Neither Tired-Energ	Energ.	Ext. Energ.
Satiation.....		2	5	5	
Non-Satiation.....		4	6	1	1

TABLE IX

SUBJECTIVE FEELING OF FATIGUE AFTER COMPLETION OF TASK (EXP. II)

	Ext. Tired	Tired	Neither Tired-Energ.	Energ.	Ext. Energ.
Satiation.....	2	5	5		
Non-Satiation.....		2	7	3	

TABLE X

TOTAL TIME, AND ESTIMATED TIME, REQUIRED TO BECOME SATIATED (EXP II) *

Satiation Group					Non-Satiation Group				
Subj.	Perf.† Time	Trials	Est.‡ Time	Diff.§ P.T.−E.T.	Subj.	Perf.† Time	Trials	Est.‡ Time	Diff.§ P.T.−E.T.
A.Z.....	18′15″	55	30′	+11′45″	M.M.	14′45″	16	15′	+ 15″
G.F.S...	25′20″	39	18′	− 7′20″	N.W.	13′10″	18	10′	− 3′10″
C.A.....	49′30″	157	30′	−19′30″	B.C.	39′30″	43	45′	+ 5′30″
J.D.....	57′25″	118	40′	−17′25″	M.K.	22′10″	37	11′	−11′10″
P.F.....	20′05″	36	10′	−10′05″	E.L.	11′30″	17	3′	− 8′30″
F.H.....	37′45″	90	15′	−22′45″	R.F.H.	21′04″	40	8′30″	−12′34″
C.R.....	25′20″	34	30′	+ 4′40″	B.E.C.	6′	12	6′20″	+ 20″
E.S.....	122′45″	389	45′	−77′45″	E.N.	17′20″	15	20′	+ 2′40″
K.W.....	37′41″	85	15′	−22′41″	J.F.A.	12′40″	12	10′	− 2′40″
A.M.S..	13′25″	27	5′	− 8′25″	B.B.	5′13″	8	8′	+ 2′47″
E.A.F..	25′	94	17′30″	− 7′30″	L.B.	3′25″	6	4′	+ 35″
J.L.....	83′10″	52	45′	−38′10″	T.	20′17″	23	25′	+ 4′43″
Mean ..	42′58″	98	25′02″	−17′53″	Mean	15′35″	20.58	13′50″	− 1′45″

* Satiation group continued until satiated; non-satiation group to five continuous successful trials.

† Performance time.

‡ Estimated time.

§ Minus quantity denotes underestimation; plus quantity: overestimation.

remains fairly stable. The fatigue is reported slightly greater for the satiates; this would be expected in light of the greater duration of activity.

Ten of 12 (83 percent) satiated subjects *underestimated* the time involved for them to become satiated. Seven of 12 (58 percent) non-satiated subjects *overestimated* the time required for them to learn the high-relief finger maze. The trend in the former is consistently toward underestimation; in the latter almost chance variation between over- and underestimation is evident.

The mean temporal activity time of the satiation group to become satiated was 42 min. 58 sec.; the average estimated time 25 min. 2 sec. The subjects in the non-satiation group required a mean of 15 min. 35 sec. to satisfy the criterion of learning; this was estimated as a mean of 13 min. 50 sec. The mean error of estimation was an underestimation of 17 min. 53 sec. for the former group; and a mean underestimation of 1 min. 45 sec. for the latter. It is clear that the averages show a large error of estimation for the satiated individuals and a small error of estimation for the non-satiates, although the interval to be estimated was in most instances larger for the satiates.

Discussion of Results

Under the foregoing experimental conditions—contrary to belief and certain previous findings—87 percent of our satiated subjects underestimated the time required for them to become satiated (pooled results). Fifty-two percent of a comparable group of non-satiated individuals *overestimated* the time required to achieve the criterion of learning. These results appeared to show some consistency, for Exp II confirmed the findings of Exp. I.

In order to be reasonably certain that an individual is satiated at a task, it is necessary to force him to persist until he rejects the task. This requires temporal periods of relatively long duration. There is thus introduced a factor of duration associated with satiation. The members of the satiation group (with a few exceptions) were called upon to estimate longer intervals than the non-satiates. Albeit the

evidence is far from conclusive, there is some indication (3) that intervals over 8 minutes are more often underestimated than overestimated. It seems improbable that the nearly consistent underestimation of our satiated subjects was a function of the difference of the interval to be judged, since all subjects under- and overestimated periods below 8 minutes (Table V) and overestimated periods as high as 39 minutes.

It is doubtful whether in a specific instance over- or underestimation is due to any one causal factor. Even when variables are isolated in controlled experiments and some consistency is possibly obtained in time estimation, it is a dynamic event due to the coördinate operation of multiple agents. Unity of organization, persistence, and need tensions have been shown to influence time estimation, and probably play some part in every judgment. Differentiated experimental approaches for securing estimations, *e.g.*, 'filled' or 'unfilled' time, duration of interval to be judged, moment of estimation, etc. add to the confusion in determining consistency of estimation. We agree with Harton (3) that greater attention should be given to individual judgments, the individual judging, and the dynamic factors involved.

Time estimation is not a precise affair. The subject seems to be at a loss when asked in retrospection to estimate an interval, and his confusion is intensified when he has no 'set' for such a judgment. The subject frequently uses the word 'about' in his estimations, and these are nearly always given in terms of whole units of time, *e.g.*, 1 min., 5 min., 10 min.; rarely are fractions of minutes or seconds given.

It is believed by certain experimenters (5, 8) that individuals who tend to wander to other tasks (*i.e.*, leave the field psychically while persisting at the satiating activity) underestimate the actual activity period; those who cannot divorce themselves from the satiating activity overestimate time. According to this theory, those of our satiated subjects who overestimated time could not develop 'creative substitutions'—their attention could not be diverted from the passing of time and thus time seemed to pass slowly. The difficulty of explaining the present findings on the basis of this theory

are threefold: (1) The non-satiated individuals who apparently were greatly engrossed in the task and did not focus on the passing of time—there thus being no need of creative mechanisms—both over- and underestimated the time involved; (2) the psychical divorcement from a satiating task ('creative substitutions') is not a unitary process, but intermittent. The subject leaves the behavioral field periodically in the struggle to persist, and no individual approaches total divorcement. It would seem under these circumstances that over- or underestimation is due to a compromise of the presence and absence of divorcement rather than an all or none affair; (3) the presence of creative mechanisms ('substitutions') in *all* of our satiated subjects (in varying degrees) does not favor a one-to-one correspondence between divorcement and underestimation.

Although our data indicates that there is a greater tendency to underestimate than overestimate time in a satiated state, the possibilities must be further explored before a theory of satiation may be attempted.

Summary

The following summary and conclusions seem to be warranted from this study:

1. Under the experimental conditions represented here 87 percent of the satiated subjects of two experiments *underestimated* the time required for them to become satiated.

2. Fifty-two percent (slightly more than chance) of a non-satiated group of subjects *overestimated* the time required for them to achieve the criterion of learning. No consistent tendency in time estimation is indicated.

3. Under a given set of conditions both underestimation and overestimation of time occur.

4. The greater underestimation than overestimation of time by the satiated subjects of these experiments does not seem to be fully explained by the theory that underestimation is characteristic of those individuals who can divorce themselves from the satiating activity; while those who cannot leave the behavioral field, overestimate time.

5. A theory of satiation based upon the belief of total consistent trends in time estimation, and in the operation of a single causative factor in judgments of time, does not seem to be warranted.

(Manuscript received January 30, 1939)

References

1. Barmack, J. E., Boredom and other factors in the physiology of mental effort, *Arch. of Psychol.*, 1937, No. 218.
2. Berman, A., An inexpensive, variable multiple-choice maze, *Amer. J. Psychol.*, 1939, **52**, 113–114.
3. Harton, J. J., *The influence of effort, persistence, and types of goal organization on the estimation of time*, Doctorate Dissert., Duke Univ., 1936 (in print).
4. Hoche, A., Langweile, *Psychol. Forsch.*, 1923, **3**, 258–271.
5. Lossagk, H., Experimenteller Beitrag zur Frage des Monotonie-Empfindens, *Ind. Psychotec.*, 1930, **7**, 101–107.
6. Swift, J. A., and McGeoch, J. A., An experimental study of the perception of filled and empty time, *J. Exp. Psychol.*, 1925, **8**, 240–249.
7. Thompson, Jr., L. H., Measuring susceptibility to monotony, *Personnel J.*, 1929, **8**, 172–196.
8. Wyatt, S., Fraser, J. A., and Stock, E. K. L., The effects of monotony in work, *Indust. Fat. Res. Bd. Rep.*, 1929, No. 56.

ACQUISITION AND EXTINCTION OF VERBAL EXPECTATIONS IN A SITUATION ANALOGOUS TO CONDITIONING [1]

BY LLOYD G. HUMPHREYS

Department of Psychology, Institute of Human Relations, Yale University

A preceding paper (**5**) was devoted to the effect of random alternation of reinforcement on the acquisition and extinction of conditioned responses. Sixty-six college students were trained to give conditioned eyelid responses to a light followed by a puff of air to the cornea. This training occurred under three different conditions, subjects being grouped according to the procedure used. Group I was given 100 percent reinforcement for 96 trials, the light always being followed by the air-puff, and then 24 extinction trials. Group II was given 50 percent reinforcement for 96 trials, the light being followed by the air-puff on only one-half of the trials, and then 24 extinction trials. Group III was given 100 percent reinforcement for 48 trials, with 48 interspersed rest intervals where non-reinforced trials occurred in Group II, followed by 24 extinction trials.

When the results obtained under the three conditions of reinforcement were analyzed, no significant difference in strength of conditioned responses was discovered. In spite of comparable acquisition performance, however, Group II (50 percent reinforcement) responded at a consistently higher level throughout extinction than Groups I and III (100 percent reinforcement). As these results conflicted with traditional conditioning theory, which posits specific increments and decrements in excitatory strength for each reinforced and non-

[1] This article is from a dissertation submitted to the Department of Psychology and the Committee on Graduate Study of Stanford University in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The writer wishes to express his indebtedness to Professor E. R. Hilgard under whose direction the work was done and to Professors C. P. Stone and R. T. Ross for assistance in the preparation of the original manuscript.

reinforced presentation of the conditioned stimulus, an alternative hypothesis was necessary. The hypothesis that the stimulating situations which the reacting organism expects, *i.e.*, is prepared for, largely determine its reactions in a conditioning situation was found to be congruent with the data presented.[2]

The present experiment was designed to throw additional light upon the functioning of expectancy uncomplicated by the 'effect' of an unconditioned stimulus. The acquisition performance of Group II (50 percent reinforcement) can be explained, for example, on the basis that considerable uncertainty concerning the appearance of a nocuous stimulus (the air-puff) can be compensated for by the strength of the avoidance reaction to that stimulus. In the words of Tolman and Brunswik (**10**), both 'reliable' and 'ambiguous' cues of a 'bad' means-object are grouped together by the organism as 'hazardous.' This factor cannot be separated from simple expectancy in the conditioning experiment, but a simplified situation analogous to conditioning can be devised in which expectancy alone is known to be operating and the results obtained from it compared with the acquisition and extinction of conditioned responses proper. The situation chosen utilized verbal expectations, *i.e.*, guesses, concerning the possible appearance of the second of two successive stimuli.

Apparatus and Procedure

On the extreme left side of a circular, vertical board 3 feet in diameter was placed a 60 watt electric light, on the extreme right another light of equal intensity. The illumination of these lights was manipulated by switches behind the board and out of sight of the subjects. The lights and their spatial disparity were easily observed by the group of subjects from various angles in the experimental room.

This apparatus was used in a way analogous to a conditioning experiment. After a ready signal the left light (the conditioned stimulus) was always turned on by the experimenter and sometimes followed after 5 seconds by the right light (the unconditioned stimulus). The subjects were instructed to guess whether or not the left light would be followed by the right light. Thus their guesses (anticipations of the right light) served as the 'conditioned responses.'

Two orders of trials were constructed to parallel as closely as possible those used for groups I and II of the conditioning experiment. For convenience these will be

[2] Expectancy as used in this hypothesis is not necessarily verbal, although purely verbal aspects are explored in the present investigation.

called series I and II, respectively. Series I had 24 'acquisition' trials, in which the right light was always 'reinforced,' *i.e.*, followed by the left light, and 12 'extinction' trials, in which the right light never followed the left light. Series II had 24 'acquisition' trials, in which the right light followed the left light only 50 percent of the time in an unpredictable order, and 12 'extinction' trials never followed by the right light. The order of 'acquisition' trials for Series II deviated from chance in that no more than two trials of a kind ever occurred together. In order to control any effects of possible symbol- or order-preferences, subjects were also asked to imagine flipping a coin 36 times and to put down an *X* to represent a head, an *O* to represent a tail, these being the symbols used, respectively, for an expected right light and for its expected absence in the experiment proper.

Subjects were 78 male and female elementary psychology students, most of whom had not served as subjects in conditioning experiments. Responses were recorded on prepared mimeographed sheets with appropriately numbered blanks corresponding to the order of trials used. All results were obtained from the group of subjects in a single session, the estimates of a chance order of the two symbols being first obtained, followed by the presentation of Series I and then of Series II.

Results

Acquisition

Figure 1 presents the acquisition results for series I and II and for the 'chance' series.[3] Percentages of 'conditioned responses,' *i.e.*, anticipations indicating that the light on the right is expected to follow the one on the left, are derived from groups of 4 trials in order to facilitate presentation and to smooth the curves.

The acquisition curve for Series I starts at approximately the 50 percent point, the expected chance score, and rises quite rapidly to almost 100 percent. The ultimate frequency reached is comparable to the frequency of conditioned eyelid responses after 96 trials.

The Series II curve starts near the 'chance' series, but there is only a slight tendency for the two to vary together thereafter. Obviously no acquisition can occur as it is just as easy to respond with an *O* as with an *X*, and both are correct equally often.

[3] Although outside the scope of this paper, it is interesting to note that at least the first five symbols in the 'chance' series were not distributed as probability theory would indicate. Twenty-three of the 78 subjects submitted the order *X*, *X*, *O*, *X*, *O* as their best guess of a random distribution. Five others submitted the order *O*, *O*, *X*, *O*, *X*. This identical order was the most frequent one in the well-known Zenith 'telepathy experiments' (1). There was, in fact, indication of extra-chance bias as far down in the list as the 20th guess.

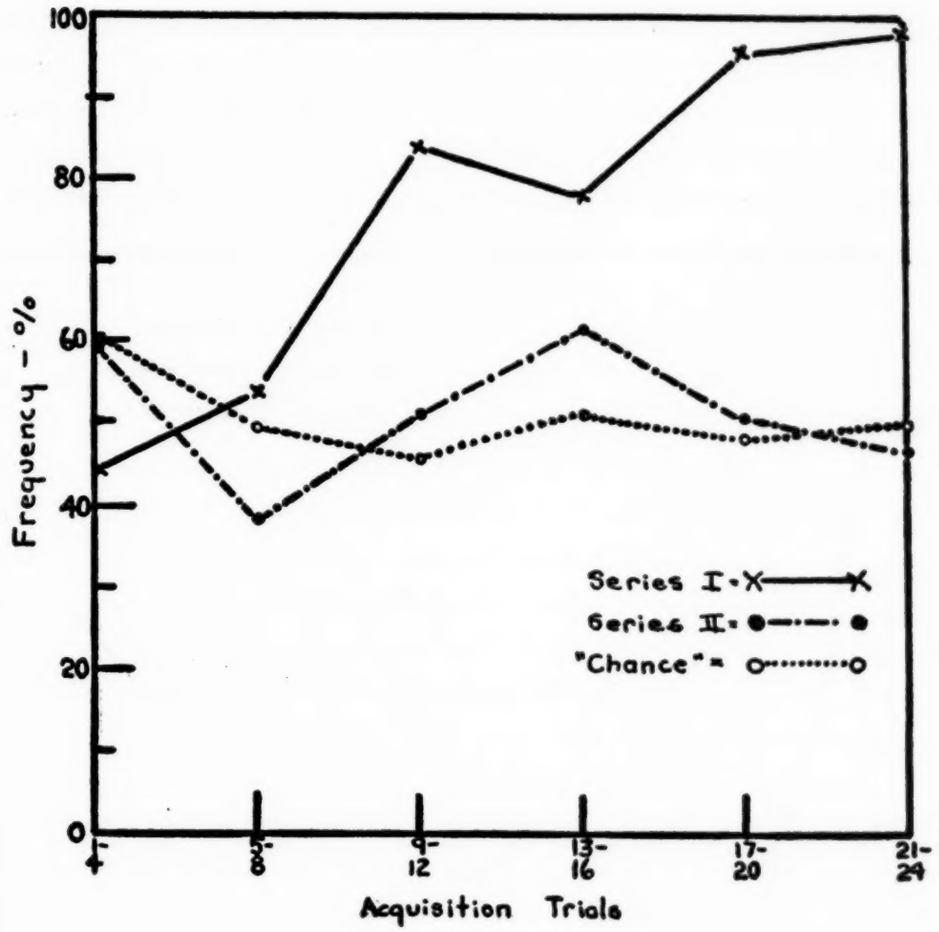

FIG. 1. Acquisition of verbal anticipations of an 'unconditioned stimulus.' The percentages refer to reported expectations of 78 subjects that a second light would appear following a first light. Series I, second light invariably appears. Series II second light appears on half of the trials. 'Chance' series is a control on order-preferences, independent of the series of lights.

Extinction

The curves by individual trials for extinction are shown in Fig. 2. Series I shows the rapid, negatively accelerated drop characteristic of conditioned responses following continuous reinforcement (5, Group I). Series II rises to the third trial, as no more than two 'non-reinforced' trials had occurred together previously, and then falls off somewhat less rapidly than Series I, this being characteristic of conditioned responses following 50 percent reinforcement (5, Group II). Only three trials are required for Series I to reach a 10 percent level of response while nine trials are necessary for Series II although there was a pre-extinction difference of approxi-

mately 50 percent in favor of Series I. It is thus more difficult to change from an expectancy of irregularity to one of regularity (Series II) than it is to change from one of regularity to an opposite regularity (Series I). This confirms the hypothesis of the preceding report (5) that the greater resistance

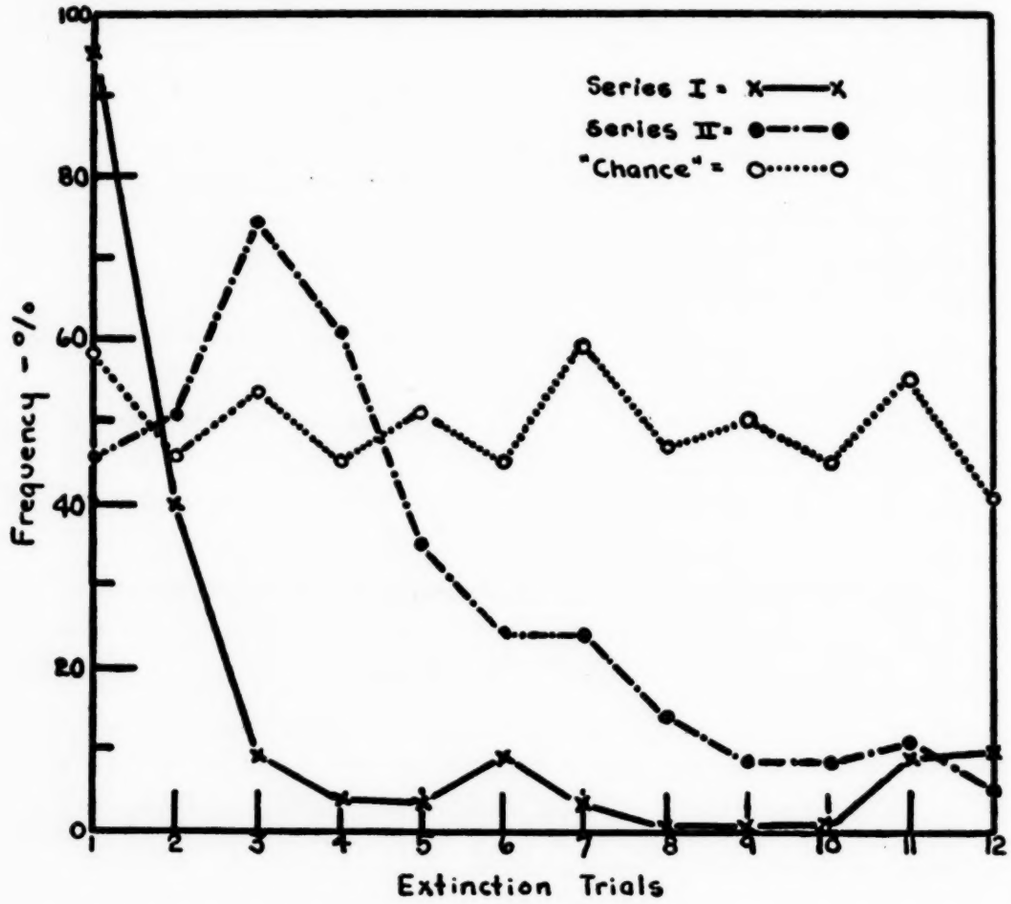

FIG. 2. Extinction of verbal anticipations of an 'unconditioned stimulus.' The percentages refer to reported expectations of 78 subjects that a second light would appear following a first light. Neither in Series I nor in Series II does the second light appear. Differences are to be attributed to the prior conditions of 'reinforcement.' 'Chance' series is a control on order-preferences, independent of the series of lights.

to extinction following 50 percent, as contrasted to 100 percent, reinforcement is due, at least in part, to greater difficulty in forming an expectancy of continuous non-reinforcement.

DISCUSSION

It has been shown that extinction of verbal expectations of a stimulus duplicates rather closely the extinction of conditioned eyelid reactions. This indicates that expectancy is,

at least, an important variable in conditioning. There are certain differences in the results, however, which indicate the presence of other variables as well.

Differences in acquisition are probably the more striking. The 50 percent reinforcement group showed as adequate conditioning as the 100 percent group in the eyelid experiment, which is considerably above the 50 percent level of expectancy in this study. Yet, under uniform conditions of 100 percent reinforcement, conditioned eyelid responses are acquired less rapidly than verbal expectations.[4]

Acquisition above expectation in the eyelid experiment is to be attributed to the presence of a noxious stimulus (the air-puff), *i.e.*, it is easier for the subject to respond than not to respond whether expectation is 50 percent or 100 percent. It accordingly follows that performance, within limits, is a product of intensity of unconditioned stimulus and expectancy. One would thus predict that significant decrement in conditioning resulting from the introduction of non-reinforced trials could be produced, contrary to the results previously reported (**5**), either by decreasing the intensity of the air-puff or by increasing the number of such non-reinforced trials.

The slower rate of acquisition of conditioned eyelid responses is not, of course, to be explained by the presence of a noxious stimulus. It may, however, be due to the instructions used. The subject is told to remain perfectly passive, neither to inhibit responses nor to provoke them, and to fixate directly ahead for the duration of the stimuli. In contrast to this type of instructions, the subject in the present experiment was told to react one way or the other and to do his best to react 'correctly.' In addition to the passivity engendered in conditioning, the instructions to fixate are actually antagonistic to the appearance of anticipatory conditioned responses.[5] If the instructions were altered to reduce

[4] This difference in rate probably does not hold for conditioned psycho-galvanic responses (**4**). The writer has made a more detailed comparison of these two responses elsewhere (**6**).

[5] Schilder (**8**, **9**) and Hilgard (**2**) have discussed the importance of conflicting tendencies in conditioning; evidence for the phenomenon has been obtained by Marquis and Porter (**7**) in the conditioning of voluntary reactions. The small effect of negative

this conflict, *e.g.*, by minimizing the set to fixate, one would predict that the rate of conditioning would be accelerated.

The difference in extinction results between the verbal expectancy and conditioning experiments is not as obvious as the acquisition differences just discussed. Although extinction is slower in both experiments following 50 percent reinforcement than 100 percent reinforcement, the discrepancy is longer maintained in conditioning; *i.e.*, the extinction curves of conditioned responses did not meet in 24 trials, while only 11 trials were required for the curves of verbal expectations. As the number of acquisition trials and pre-extinction level of response differed in the two experiments, and as Series II followed Series I in the experimental procedure, a conclusion concerning differential extinction rates is uncertain. The evidence does not, at any rate, preclude the presence of factors in extinction additional to expectancy.

Summary and Conclusions

To better estimate the possible influence of expectancy of stimulating situations in a conditioning experiment, the present experiment was designed to study simple expectancy uncomplicated by reactions to an unconditioned stimulus. Responses were obtained from a group experiment with 78 college students. Subjects were presented with two lights: a 'conditioned stimulus' light and an 'unconditioned stimulus' light. The subjects were instructed to write down on prepared blanks, in two series of 24 acquisition and 12 extinction trials, whether or not they expected the 'unconditioned stimulus' light to be turned on after the 'conditioned stimulus' light appeared. Thus their verbal anticipations of the 'unconditioned stimulus' light served as the 'conditioned' responses.

Two series of presentation orders, designated I and II, were composed, analogous to the situations of 100 percent and 50 percent reinforcement used in a previous conditioning experiment. When a comparison is made between the con-

voluntary instructions in the experiment of Hilgard and Humphreys (3) may have been due to the presence of similar, self-induced instructions in their control group.

ditioning and the verbal expectancy experiments, the following conclusions can be drawn:

1. Series I shows acquisition of the verbal anticipations comparable to the frequency of conditioned responses under 100 percent reflex reinforcement. The former type of response is, however, acquired more rapidly.

2. The extinction curve for Series I shows the rapid, negatively accelerated drop characteristic of extinction following 100 percent reflex reinforcement.

3. The extinction curve for Series II shows a much slower drop than Series I, this being characteristic of extinction following 50 percent reflex reinforcement.

4. Conclusions 2 and 3 confirm the hypothesis that the greater resistance to extinction following 50 percent, as contrasted to 100 percent, reflex reinforcement is due, in part at least, to greater difficulty in forming an expectancy of continuous non-reinforcement.

(Manuscript received March 8, 1939)

References

1. Goodfellow, L. D., A psychological interpretation of the results of the Zenith radio experiments, *J. Exp. Psychol.*, 1938, **23**, 601–632.
2. Hilgard, E. R., An algebraic analysis of conditioned discrimination in man, *Psychol. Rev.*, 1938, **45**, 472–496.
3. Hilgard, E. R. and Humphreys, L. G., The effect of supporting and antagonistic voluntary instructions on conditioned discrimination, *J. Exp. Psychol.*, 1938, **22**, 291–304.
4. Hovland, C. I., The generalization of conditioned responses. IV. The effects of varying amounts of reinforcement upon the degree of generalization of conditioned responses, *J. Exp. Psychol.*, 1937, **21**, 261–276.
5. Humphreys, L. G., The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions, *J. Exp. Psychol.* (in press).
6. Humphreys, L. G., Extinction of conditioned psycho-galvanic responses following two conditions of reinforcement (to appear).
7. Marquis, D. G. and Porter, J. M., Differential characteristics of conditioned eyelid responses established by reflex and voluntary reinforcement, *J. Exp. Psychol.*, 1939, **24** (in press).
8. Schilder, P., The somatic basis of the neurosis, *J. Nerv. Ment. Dis.*, 1929, **70**, 502–519.
9. Schilder, P., Conditioned reflexes, *Arch. Neur. Psychiat.*, 1929, **22**, 425–443.
10. Tolman, E. C. and Brunswik, E., The organism and the causal texture of the environment, *Psychol. Rev.*, 1935, **42**, 43–77.

STUDIES IN THERMAL SENSITIVITY: 11. EFFECTS OF STIMULATOR SIZE IN SERIATIM COLD-MAPPING

BY WILLIAM LEROY JENKINS
Lehigh University

In a previous article,[1] it was shown that highly reliable results can be obtained from untrained subjects in seriatim cold-mapping with 2 mm and 3 mm diam. stimulators. In the present study, a wider range of stimulator sizes has been used upon a larger number of untrained subjects, to secure some preliminary data upon the influence of stimulus area.

Procedure.—Briefly, the seriatim method as here carried out involves repeated stimulation of a suitable checkerboard pattern so that a 'score' is obtained by summing the reports. Four categories are employed: 0 for neutral, 1 for weak, 2 for medium and 3 for strong cold. (Warm, if experienced, is simply reported as 'W'.) Six rounds of stimulation constitute a series. Thus the maximum score is 18 and the minimum 0. The scores are merely pseudoquantitative, since the categories 1, 2 and 3 are ordinal and not cardinal numbers. Defense of the statistical procedure and an analysis of the probable artifacts in mapping have been presented in the previous paper.[2]

All of the experiments here reported were performed by the group method, three students to a group, one acting as experimenter and the other two alternating as subject and recorder. To equalize the effects of adaptation as far as possible, a uniform time of 3½ min. per round was maintained; so that approximately 7 min. elapsed between successive stimulations of the same squares. A practically random order of stimulation was employed, and the general sequence continually shifted so that the subject (who worked with his eyes closed) could not memorize a pattern of reports.

Five stimulator sizes were used: 3 × 3 mm, 3 × 1½ mm, 2 × 2 mm, 1½ × 1½, and a circular stimulator 1 mm in diam. The last-named approaches the limit for accurate manual placement.

RESULTS

Because the same students were not available for all of the experiments, the following results are not strictly comparable on a quantitative basis. However, the major implications of the data are sufficiently clear.

[1] Jenkins, W. L., Studies in thermal sensitivity: 9. The reliability of seriatim cold-mapping with untrained subjects. *J. Exper. Psychol.*, 1939, *24*, 278–293.

[2] *Op. cit.*, 282–284.

1. *Reliability Coefficients.*—Table 1 summarizes the reliability coefficients obtained when the scores from two successive series of 6 rounds each are compared. With all stimulator sizes, at least 75 percent of the subjects reached .70 and approximately half of them .80—with a few exceptional individuals making .90 or higher. Where low reliability was obtained, the chief cause seemed to be the failure to discriminate a sufficient range of experiences. The few subjects below .60 typically employed only two categories of report. Occasionally, a subject who could do well with the largest size (3 × 3 mm) would fail miserably with the smaller ones.

TABLE 1

RELIABILITY COEFFICIENTS *

Stim. Size	Below .60	.60 .64	.65 .69	.70 .74	.75 .79	.80 .84	.85 .89	.90 .94	.95 1.00	Total Cases
3×3.....	0	2	3	5	2	7	8	3	1	*31*
3×1½.....	1	0	1	3	5	3	6	1	0	*20*
2×2.....	11	4	6	8	14	13	20	6	0	*82*
1½×1½....	2	0	3	2	2	5	6	1	0	*21*
1 mm.....	0	2	3	2	6	2	4	5	4	*28*

* Figures in the table indicate the number of cases within the given step interval.

Virtually all subjects attained somewhat better results with the 3 × 3 mm than with the 2 × 2 mm stimulator.

Possibly the most surprising result is the high percentage of coefficients over .90 with the 1 mm diam. It might be thought that this was due to reporting on an all-or-none basis. Such was not the case. Although the '3' category virtually disappears at this level and the number of consistent zeros is large, 9 of the 11 subjects above .90 used both '1' and '2' consistently. It will be recalled that the 1 mm diam. is the size of stimulator used by Dallenbach in his study of maps repeated on four different days.[3] Since some of our *untrained* subjects were able to reach .90 or better, it is reasonable to assume that Dallenbach's *trained* observers would do equally well. This supports our previous contention [4] that the discrepancies in the four mappings were indications of actual gradations of

[3] Dallenbach, K. M., The temperature spots and end-organs. *Amer. J. Psychol.*, 1927, **39**, 402–427.

[4] *Op. cit.*, 280.

sensitivity within the area (hills and valleys) and not of poor work on the part of experimenters and observers.

In the balance of this article, only the data from subjects with reliability coefficients of .80 and over will be cited. This criterion is probably unnecessarily stringent, but will avoid any possibility that the conclusions are distorted by purely random reporting.

2. *Distributions of Scores.*—Table 2 shows the *average* distribution of scores for each of the five stimulator sizes, to-

TABLE 2

DISTRIBUTIONS OF SCORES

Stim. Size		0–1	2–3–4	5–6–7	8–9–10	11–12–13	14–15–16	17–18
3×3	(Av.)....	7%	13%	21%	20%	18%	15%	6%
	(Min.)...	24	22	20	14	10	6	4
	(Max.)..	0	2	10	18	28	30	12
3×1½	(Av.)....	12%	18%	18%	16%	16%	11%	9%
	(Min.)...	48	26	26	0	0	0	0
	(Max.)..	0	0	4	6	16	32	42
2×2	(Av.)....	13%	16%	23%	19%	15%	9%	5%
	(Min.)...	66	22	8	0	4	0	0
	(Max.)..	0	0	4	20	22	26	28
1½×1½	(Av.)....	15%	25%	25%	15%	11%	7%	2%
	(Min.)...	47	25	28	0	0	0	0
	(Max.)..	2	3	12	20	26	27	10
1 mm	(Av.)....	59%	16%	12%	5%	5%	2%	1%
	(Min.)...	80	5	15	0	0	0	0
	(Max.)..	22	20	10	15	20	8	5

gether with the individual distributions for the subjects having the maximum and minimum gross scores in each case. Only subjects with reliability coefficients of .80 and over are included. The grouping: 0–1, 2–3–4, 5–6–7, etc. represents roughly the four category levels and the three intermediate positions.

This *average* distribution of scores is a purely mathematical construct. The actual distributions varied between the maximum and minimum levels shown, so that the average can scarcely be considered even as typical. However, two chief points do stand out: (1) With stimulators 1½ × 1½ and larger, the distributions are all very similar and there is con-

siderable overlapping. (2) Between the 1½ × 1½ and 1 mm diam. sizes, there appears to be a sharp break which is likewise reflected in the maximum and minimum scores.

3. *Average Scores.*—Table 3 shows the average scores for subjects with reliability coefficients of .80 and over, and the mean score for each stimulator size. The median score for each size agrees closely with the mean, except in the case of the 1 mm diam. stimulator where a few high values cause an

TABLE 3

AVERAGE SCORES *

Average Scores	Stimulator Sizes and Areas				
	3 × 3 9.0 mm.²	3 × 1½ 4.5 mm.²	2 × 2 4.0 mm.²	1½ × 1½ 2.25 mm.²	1 mm. 0.8 mm.²
0–1	—	—	—	—	—
1–2	—	—	1	—	8
2–3	—	1	1	1	1
3–4	—	—	2	—	2
4–5	—	1	3	3	2
5–6	1	—	6	1	1
6–7	4	1	3	2	1
7–8	1	2	5	3	—
8–9	5	1	4	1	—
9–10	2	2	4	—	—
10–11	2	1	7	—	—
11–12	2	—	1	1	—
12–13	2	—	—	—	—
13–14	—	—	2	—	—
14–15	—	1	—	—	—
Total	19	10	39	12	15
Mean score for group	*9.0*	*8.1*	*7.7*	*6.4*	*2.8*

* Figures in the table indicate number of cases within the given step intervals.

upward distortion of the mean. Taking this into account, it appears that the averages for the 1 mm diam. and 1½ × 1½ mm sizes are closely proportional to the relative stimulus areas. Taken at its face value, this would indicate a roughly linear summation effect within this range. Above this level, however, a plotted curve would flatten out rapidly. A change from 4.5 mm² to 9.0 mm², for example, brings about an increase of less than one unit in the average score. Again, it must be emphasized that these 'average' values are not quantitatively significant but may be taken only as general indicators of a trend.

4. *Maps.*—The general appearance of the maps derived from subjects showing reliability coefficients of .80 and over does not differ significantly from those shown in the earlier study.[5] With all stimulator sizes, there is a tendency for similar levels of sensitivity to be grouped. Even with the 1 mm diam., the picture is typically one of hills and valleys, like a topographic map, and not of a random arrangement of scattered 'cold spots.' As indicated by the wide range of average scores, the general level of reports varies markedly from subject to subject, and the appearance of the maps changes correspondingly.

Discussion

The foregoing should be considered as no more than a preliminary attack on the problem. It is evident that simple linear summation does not occur with all stimulator sizes. It may be an oversimplification to assume that it holds even in the lowest range. The chief value of the study lies in showing the feasibility of the seriatim technique for this purpose. Since good subjects can report consistently over a range of stimulus sizes, as shown by the high reliability coefficients, it should now be possible to map the same checkerboard area with two different sizes. The comparative scores will then give a basis for evaluating summation effects with supra-liminal stimulation.

Summary

Seriatim cold-mapping with a range of five stimulator sizes, the smallest 1 mm diam., gives satisfactory reliability coefficients (.80 or over) with all sizes in the majority of untrained subjects. Average scores indicate roughly a linear summation effect at the lowest levels, but the curve flattens out rapidly with the larger sizes. Even with the 1 mm diam. stimulator, the maps tend to show grouping of similar scores into hills and valleys, and not a random arrangement of separate cold spots. By mapping the same area with two different sizes of stimulator, it should be possible to attack the problem of summation with supra-liminal stimuli.

(Manuscript received February 17, 1939)

[5] *Op. cit.*, 290.

STUDIES IN HIGH SPEED CONTINUOUS WORK. II. DECREMENT

BY B. R. PHILIP
Queen's University

Most work curves studied, for example, those of Ioteyko (1), Neifeld and Poffenberger (2), and Flügel (3), show a pronounced decrement almost from the beginning of the work period, which permits the fitting of parabolic or hyperbolic equations to them. A previous study (4) of curves of work done at high speed for 6 or 7 hours showed a characteristic periodicity sufficiently regular to satisfy statistical criteria. The marked variability of the curves, especially when short time intervals are used, and their pronounced individuality of pattern, preclude the fitting of simple equations to them. To avoid these difficulties another experiment was devised in which Ss work for repeated periods of moderate length, followed by short rests. By averaging the same phases of each short run the decrement is brought out more effectively. This we shall call Experiment B to distinguish it from the one previously reported, henceforth called Experiment A.

The Ss of Experiment B were 2 college men who performed the tapping under the same conditions as in Experiment A, for approximately the same length of time (6 hours). However Ss did not work continuously; they went through a series of 28 ten min. runs of continuous tapping at maximum speed, with one min. rest between runs. Breakdown of the apparatus spoiled the records of 2 of the runs of one S, so that his results are given for 26 runs only. For 2 mins. after every rest the results were registered at 10 sec. intervals. Following the initial 2 mins., readings were taken every 30 secs. for the remaining 8 mins. of the run. The 6 hour experiment was finished off with a half hour continuous run prior to which, instead of the customary one min. rest, a rest of 5 mins. was taken. No minimum rate, based upon a pre-experimental average, was set, but Ss were urged repeatedly to work at maximum possible speed. The work started for both Ss at approximately 3.00 P.M. and lasted till 9.00 P.M.

In Experiment A, 12 Ss worked continuously for about 6 hours, tapping at maximum speed. No rests were given except those that occurred due to a breakdown of the apparatus, or after S had reached a low point, not less than 80 percent of a pre-experimental test of 10 mins.

Results

Philpott (5) has suggested two extreme forms of the output curve. "Type A is that of a single homogeneous task within which no threads are engendered save at the nominal moment of starting work. It will presumably be of marked hyperbolic shape, with geometric periods and increasing amplitudes. Apparent fatigue will be maximal. Type B is of a succession of relatively short operations. The resulting curve will obviously be a succession of short hyperbolas joined end to end. To the extent that all threads change with the commencement of each short operation, the resulting curve will be horizontal with reference to the base line." Experiment A sets a task which in general conforms to Philpott's Type A; the results, however, do not quite conform to his analysis. Experiment B does yield a series of short hyperbolas joined end to end, and to that extent fits his analysis quite well. When, however, the course of the entire period of work is portrayed by plotting the maxima, the average, and the minima of each run, a curve is found whose characteristics closely resemble the previously described curves of Experiment A. The amount and rate of work done in this experiment is also comparable with that of Experiment A, as may be seen from Table I. The curves are plotted in

TABLE I

Amount of Work Done

SS	Test Rate	Total Time	Total Number of Taps	Range	σ Dist.	Time of Maximum
S1.....	251.9	310	75,158	202–321	6.15	100
S2.....	221.1	310	68,096	186–278	6.65	100

Fig. 2. To be noted are the initial spurts, the initial slump, the absence of any marked decline in efficiency, the periodic oscillations, the deferred occurrence of the absolute maxima, whether of single mins. or of 10 min. averages, the greater variability of the maxima as contrasted with the minima, and the end spurt. Obviously any pronounced decrement, such as are found in the curves of Ioteyko and Neifeld and Poffen-

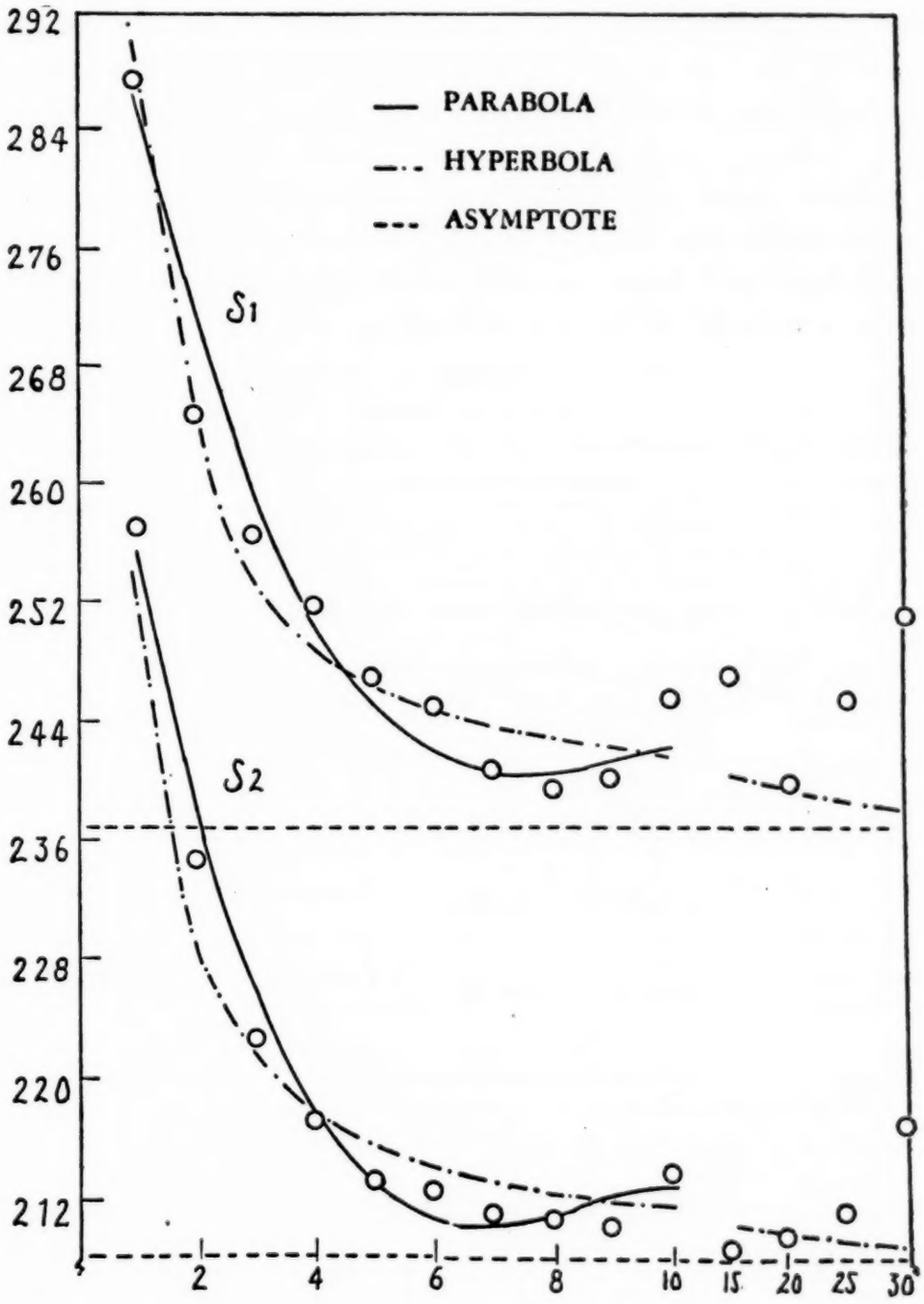

FIG. 1. Work decrement minute by minute. Ordinates, number of taps per min. Abscissæ, minutes. A more restricted time scale is used for final 20 min. period.

berger, must be sought, not in the general trend of the entire work period, but rather in the course of the short runs.

Figure 1 shows the curves obtained by averaging the short runs min. by min. On each curve are fitted the parabolas and hyperbolas obtained by the method of least squares,

whose equations and criteria of goodness of fit are recorded in Table II. It is seen that both curves fit extraordinarily well, with the third degree parabola giving slightly the better fit.

When these theoretical curves are applied (Fig. 2*C*) to the 30 min. run at the end of the working period—data which had not been used in their computation, the third degree parabola gives unsatisfactory values. On the other

TABLE II
DATA ON CURVES

Curve	Third Degree Parabola	Asymptote	χ^2	P	n
S1	$Y=306.628-22.578X+2.5025X^2-0.0889X^3$		0.3000	99	10
S2	$Y=278.541-25.838X+3.1770X^2-0.1250X^3$		0.1726	99	10
Initial Spurt	$Y=283.150-0.400X-1.689X^2+0.1430X^3$		0.1366	99	10
30 Min. Average	$Y=258.520-0.735X+0.0077X^2-0.0095X^3$		0.3695	99	11
	Hyperbola	Asymptote			
S1	$Y=\frac{X-5}{0.004218X-0.000810}$	$Y=237.08$	0.3622	99	10
S2	$Y=\frac{X-5}{0.004807X-0.000866}$	$Y=208.08$	0.3392	99	10
Initial Spurt	$Y=\frac{X-5}{0.004141X-0.001091}$	$Y=243.90$	0.8407 7.7440	99 56	9* 10
30 Min. Average	$Y=\frac{X-6}{0.004160X-0.000710}$	$Y=240.39$	0.1964 4.0814	99 94	10* 11

* Value of χ^2 and P when the first interval is omitted.

hand the hyperbola fits quite closely, and the asymptotes to the Y axis ($Y = 237.08$ for S1, and $Y = 208.10$ for S2) predict the decrement trend. Thus, both curves are good mathematical expressions of the data from which they have been derived, but the parabola is useless for prediction purposes, while the hyperbola predicts results beyond the data from which it has been computed.

The cessation of the runs at the 10th. min. brings on the end spurt in which the scores of the last min. or two are increased. A more accurate picture of the initial slump is obtained by discarding these end intervals. The hyperbolas

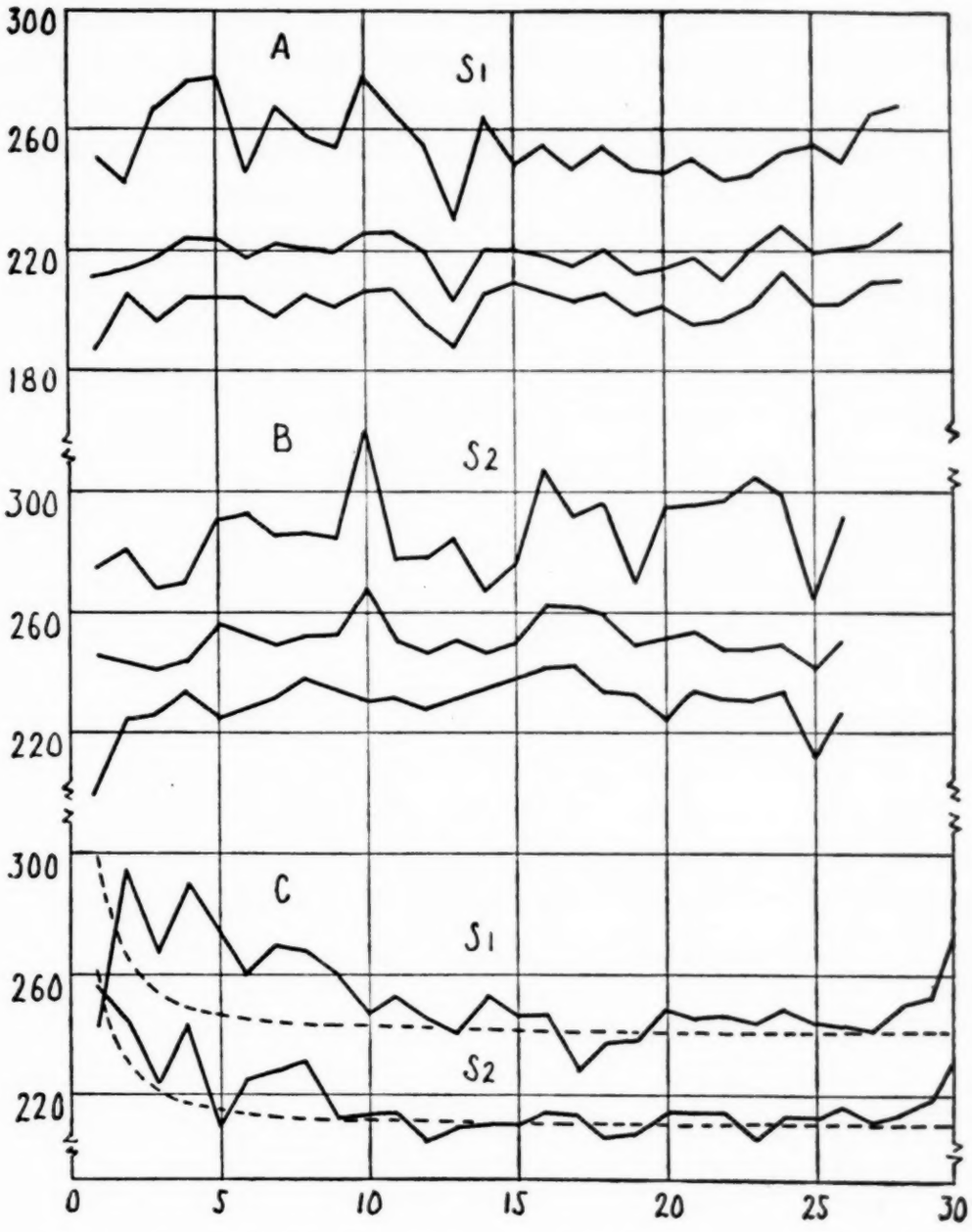

FIG. 2. *A* and *B*. Maxima, averages, and minima (upper, middle and lower values), of succession of 10 min. runs for S1 and S2. Ordinates, number of taps per min. Abscissæ, number of 10 min. run. *C*. Fitting of parabolas (dotted lines), to 30 min. final period. Ordinates, number of taps per min. Abscissæ, minutes.

recomputed on the basis of 8 intervals for S1 and 9 intervals for S2 have their equations changed slightly, and the corrected asymptotes ($Y = 234.4$ for S1 and $Y = 206.7$ for S2) fit the trend of the continued work.

So close was the fit of the calculated curves to the pooled data of the short runs, and so accurately did the asymptotes

predict the trend of the working rate when the time is extended considerably beyond that from which the calculations of the curves were derived, that it suggested the prediction of the trend of the work of Experiment A from the pooled values of the first 10 mins. where the decrement is most apparent.

This is shown in Fig. 3*A*. It will be noted from the curve

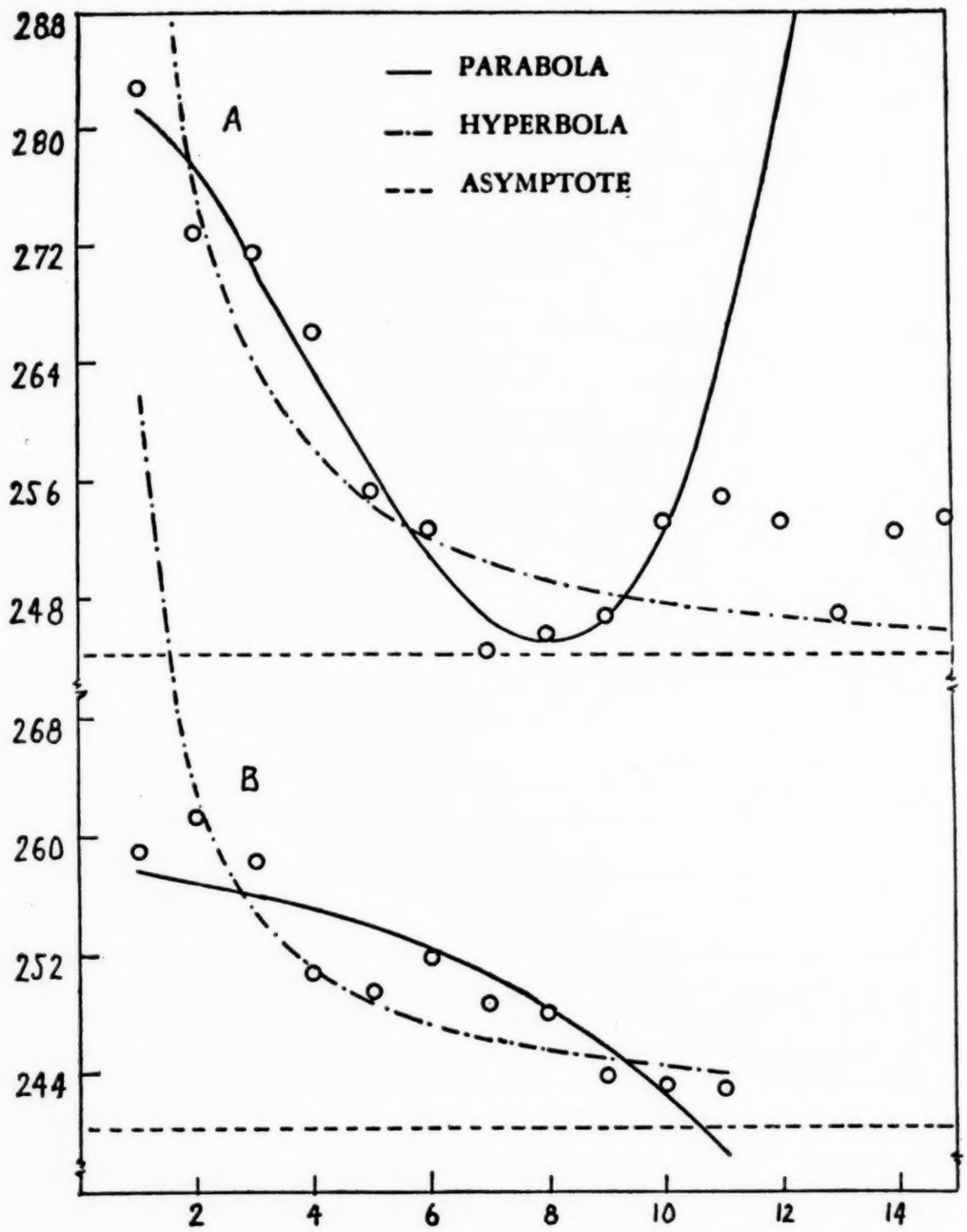

Fig. 3. *A*. Fitting of parabola and hyperbola to initial slump (averages of 12 Ss). *B*. Fitting of parabola and hyperbola to 30 min. averages (12 Ss). Ordinates, number of taps per min. Abscissæ, Curve *A*, minutes; Curve *B*, 30 min. periods.

and from Table 2 that both types of curves fit the data, except for the initial point of the hyperbola. For points beyond the 10th. min. *i.e.* beyond the data from which the curves are calculated, once again the parabola gives unsatisfactory values, while the hyperbola is a fair fit. When the hyperbola is calculated for the first 8 mins. only (viz. to the end of the initial slump), the parameters are somewhat changed, and the asymptote becomes $Y = 237.0$.

The marked variability of the work estimated in intervals of 5 mins. or less is diminished by increasing the size of the interval. Accordingly a curve was drawn from averages based on 30 min. intervals, and carried to 11 intervals, *i.e.* $5\frac{1}{2}$ hours. A slight adjustment (about 1 percent) had to be made for the last two intervals, as 2 Ss had finished their work at the 9th. interval. The curve is shown in Fig. 3*B*. A third degree parabola and an hyperbola, whose equations and goodness of fit are given in Table II, were fitted to this curve. With the exception of the first interval for the hyperbola both agree well, though in this case a straight line might apply with fair results. The asymptote of the hyperbola ($Y = 240.4$) agrees closely with that previously obtained from the initial 10 min. curve ($Y = 243.9$), and that of the initial slump ($Y = 237.0$). In the case of the pooled results at least, the trend of the work continued for $5\frac{1}{2}$ hours is adequately represented by the asymptote of the hyperbola fitted to the scores of the initial slump. Up to this point no marked reduction in working rate had taken place: had Ss become exhausted the efficiency would undoubtedly have dropped well below this limiting value.

The pooling of the 30 min. averages of the 12 Ss smooths out the curve and hides the characteristic individual pattern. The adequacy of the asymptote fitted to the initial slump in predicting the results hours ahead does not of course necessarily mean that this relationship holds in every individual case.

An attempt was made to determine the relationship in the individual case. As has been stated, so distinct are the work patterns that it is impossible to fit hyperbolas to all

curves to compare their asymptotes with those of the initial slump. Accordingly, to indicate the trend after several hours of work the 30 min. average at the 9th. interval was selected. This interval was chosen because values beyond it are missing in 2 cases (though some Ss went up to the 14th. interval), and the variability increases greatly in the last few intervals. The comparison between these 30 min. averages and the asymptotes of the hyperbolas fitted to the initial slump is shown in Table III. There is an agreement within 3 percent

TABLE III

COMPARISON OF THE RATE OF 9TH INTERVAL WITH THE ASYMPTOTE OF INITIAL SLUMP

Ss	1	2	3	4	5	6	7	8	9	10	11	12
9th Int.....	269	250	271	205	222	213	272	280	238	213	250	244
Asymptote.....	272	242	230	193	213	202	225	195	242	212	263	239

on the average when the results of 3 Ss are omitted. Ss who do not show this agreement have distinct patterns to their work curves: that of S3 remains level throughout the whole test; that of S7, after the 5th. interval, reaches and maintains its highest level; that of S8 has its lowest point of the whole test in the first 10 mins.

On the whole therefore, even for the individual, a fair prediction of the trend of hours of continuous work, before S's efficiency decreases rapidly due to fatigue, is obtained from the asymptote to the X axis of the hyperbola fitted to the slump during the first 10 mins.

It is of interest to compare the equations of this experiment with those of Ioteyko and Neifeld and Poffenberger. The former obtained a curve of the form $Y = a - bx + cx^2 - dx^3$, and the latter, one of the form $Y = a + bx - cx^2 + dx^3$. The parabolas here obtained differ in still other signs. Obviously the interpretation placed by Ioteyko on the signs of the parabola have no especial significance, and the formulae are merely mathematical expressions of the data, as was pointed out by Neifeld and Poffenberger.

Since the hyperbola fits the data about as well as the parabola, and in addition enables us to predict the course of

the work beyond the data from which the curves were computed, it is suggested that this mathematical form is the more satisfactory one to use as an expression of the trend of continuous work.

(Manuscript received January 26, 1939)

REFERENCES

1. IOTEYKO, J., *Les lois de l'ergographie*, Paris, 1904.
2. NEIFELD, M. R. AND POFFENBERGER, A., A mathematical analysis of work curves, *J. Gener. Psych.*, 1, 1928, 448–458.
3. FLÜGEL, J. C., Practice, fatigue and oscillation, *Brit. J. Psych.*, Monograph Supp., 1928, 13, 1–80.
4. PHILIP, B. R., Studies in high speed continuous work: I. Periodicity, *J. Exp. Psychol.*, 1939, 24, 499–510.
5. PHILPOTT, S. J. F., Meaning of G and other factors in terms of rates of output, *Brit. J. Psych.*, 1936–7, 27, 196–225.

TRANSFER OF TRAINING IN MIRROR TRACING

BY BEATRICE E. HALL

From the Psychological Laboratory of Acadia University

The many experiments on cross-education and on transfer of training have to date included few studies that permit comparison between the principles governing the two types of transfer. The present investigation is an initial attempt to gather data for such a comparison, using the mirror-maze and the general procedure followed by Cook (1) in his studies of cross-education, but always tracing with the right hand and varying the position of the mirror from practice to transfer trials. The only changes in the apparatus were the replacement of the adjustable screen by a 12 × 18 inch piece of heavy cardboard fixed in a horizontal position (on 4 wooden posts) 7 inches above the surface of the maze, and wooden stops nailed to 3 sides of the maze to permit quick and accurate placing of the mirror on the right or left side, as well as in a central position directly in front of, but on the opposite side of the maze from the subject.

The subjects were 60 students of Acadia University. These were divided into 6 groups of 10 subjects each (4 girls, 6 boys) in order to use all combinations of the temporal order of the three mirror positions—R (right) C (center) L (left), RLC, CLR, CRL, LRC, LCR. Each subject traced the maze 30 consecutive trials with 20 seconds between each trial. The position of the mirror was changed every 10 trials. Thus a subject traced the maze 10 trials with the mirror in position 1, 10 trials with the mirror in position 2, and 10 trials with the mirror in position 3. Averaging the results of the 6 groups equates differences in difficulty of right, left, and center mirror positions, since (for example) the mirror was on the right side with Groups I and II in the first 10 trials, with Groups IV and V on the second 10 trials, and with Groups III and VI in the last 10 trials. Two stop-

clocks were used for timing, one clock being stopped and the other started at each change from tracing to rest interval and vice versa. The trial by trial averages for all 60 subjects for each of the 3 mirror positions are given in Table I.

On examination the data seem to warrant certain conclusions.

1. Some transfer is apparent for at least the first 7 trials with position 2 and the first 5 trials with position 3.

2. The apparent decrease in efficiency in trials 8 to 10 with position 2 (relative to the same trials with position 1),

TABLE I

Trial	Time in Seconds		
	Mirror Position 1	Mirror Position 2	Mirror Position 3
1	46.9	32.8	29.9
2	34.2	29.8	22.9
3	29.8	26.2	22.5
4	23.7	20.8	19.6
5	20.5	19.0	19.3
6	19.2	17.4	19.1
7	18.9	16.9	17.8
8	16.9	16.1	17.7
9	15.6	16.7	17.1
10	15.3	16.9	17.6
Average.....	24.1	21.2	20.3

and trials 5 to 10 with position 3, may not contra-indicate permanence of transfer. It is possible that such permanence is masked by fatigue. A control experiment with longer rest intervals should be performed.

3. The amount of transfer is not so large as in cross-education. The parallel experiment in Cook's work shows 76 percent transfer from 'practice' to '10 transfer' trials, and 65 percent from the latter to the '20 transfer' trials (**2**, Experiment B). In the present experiment the transfer from position 1 to position 2 is 45 percent, and from position 2 to position 3, 18 percent of the amount of improvement in the preceding 10 trials.

Summary

Transfer from 10 trials mirror-tracing a star-shaped maze with the mirror in one position (right, left, center) to a second

or third position was tested with 6 groups of 10 subjects each, using all possible combinations of temporal order of right, left, and center mirror-positions. The average transfer with the 6 groups is considerable less than in cross-education experiments: 45 percent from position 1 to position 2, and 18 percent transfer from position 2 to position 3.

(Manuscript received February 6, 1939)

References

1. Cook, T. W., Studies in cross education. 1. Mirror tracing the star-shaped maze, *J. Exp. Psychol.*, 1933, **16**, 144–160.
2. ——, Further experiments in mirror tracing the star-shaped maze, *J. Exp. Psychol.*, 1933, **16**, 679–700.

AN INEXPENSIVE VOICE RELAY FOR USE WITH THE ELECTRONIC CHRONOSCOPE

BY FERDINAND HAMBURGER, JR.
The Electrical Engineering Department of The Johns Hopkins University

Since the development of the electronic chronoscope described by Dorcus and Hamburger (1), several investigators have raised a question regarding the use of this chronoscope for the measurement of voice reaction. The problem of adapting the electronic chronoscope to this measurement is readily solved by a modification of the circuit of De Silva (2) proposed by him in connection with the chronoscope which he developed.

The modification of De Silva's arrangement, referred to above, is a direct consequence of recent improvements in electron tubes and associated apparatus and results in an exceedingly satisfactory but much less expensive voice relay. This improvement consists essentially of substituting a cold-cathode gas-triode [1] for the thyratron of De Silva. A dual saving results, the cost of the gas-triode is about one-tenth that of the thyratron and in addition the cold cathode eliminates the necessity for filament heating and hence does away with two filament transformers. The resulting voice relay will operate from three 45 volt '*B*' batteries, and four 1.5 volt dry cells, and in combination with the electronic chronoscope gives a compact, portable, self-contained unit for voice reaction measurement.

The circuit using two cold-cathode gas-triodes is shown in Fig. 1. It consists simply of a resistor '*R*' and a potentiometer '*P*' connected in series across the direct current source as indicated, two relays, two cold-cathode gas-triodes, two microphones, four batteries and two microphone transformers and six resistors. In operation the potentiometer is set so

[1] Detailed descriptions of all components are given under the diagram of connections, Fig. 1.

that each gas-triode just does not discharge. An impulse from either microphone will then add sufficient voltage to the respective starter-anode circuit to cause the tube to discharge and the corresponding relay to operate. When used with the electronic chronoscope (1), the contacts of the relay '*E*' are connected to the *K*-1 terminals and those of relay *S* to the *K*-2 terminals (or vice-versa). Since the operation of this voice relay is by a make-break sequence it can be used

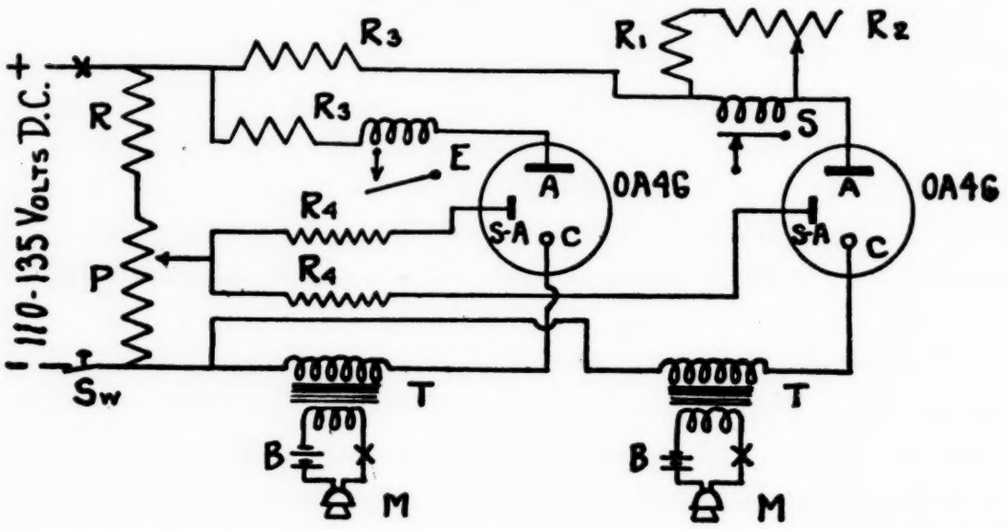

FIG. 1. Wiring diagram voice relay. *Sw*—switch, single pole, single throw; *P*—potentiometer 25,000 ohms; *R*—resistor 5,000 ohms, 1 watt; R_1—resistor 2,000 ohms, 1 watt; R_2—potentiometer 2,500 ohms; R_3—resistor 2,500 ohms, 1 watt; R_4—resistor 50,000 ohms, ½ watt; *B*—2 No. 6 dry cells; *M*—single button carbon microphone, American Microphone Co.; *T*—microphone to grid transformer Kenyon, Type R-55; *E*, *S*—single pole double throw relays, Type 3-A, 2,000 ohms, Sigma Instruments, Inc.; *OA4G*—cold-cathode gas-triode; sockets—2 octal sockets for tubes, 2 5-prong sockets for relays. Switches should be provided at the points marked *X* to de-energize the circuits when not in use.

with any other chronoscope which operates on this same sequence. It can also be arranged for other modes of operation since the relays used are of single-pole double-throw construction.

In the circuit actually constructed the two relays used are identical single-pole double-throw relays. It is obvious that for satisfactory operation the operating time of each relay must be the same. Since for identical relays, when one is used as a make relay and the other used as a break relay, the time to break is, in general, shorter than the time to make,

some adjustment of the time is necessary. Therefore, one relay, for the Experimenter, is arranged to make contact when the Experimenter microphone is actuated and this relay is adjusted to operate as quickly as possible by setting its armature quite close to the pole piece. The other relay, for the Subject, is arranged to break contact when the Subject microphone is actuated and this relay is adjusted to operate more slowly than the *E*, by setting its armature further from the pole than that of the *E* relay. Proper adjustment assures that the *E* relay closes at the same instant that the *S* opens for simultaneous actuation of the microphones. In order to

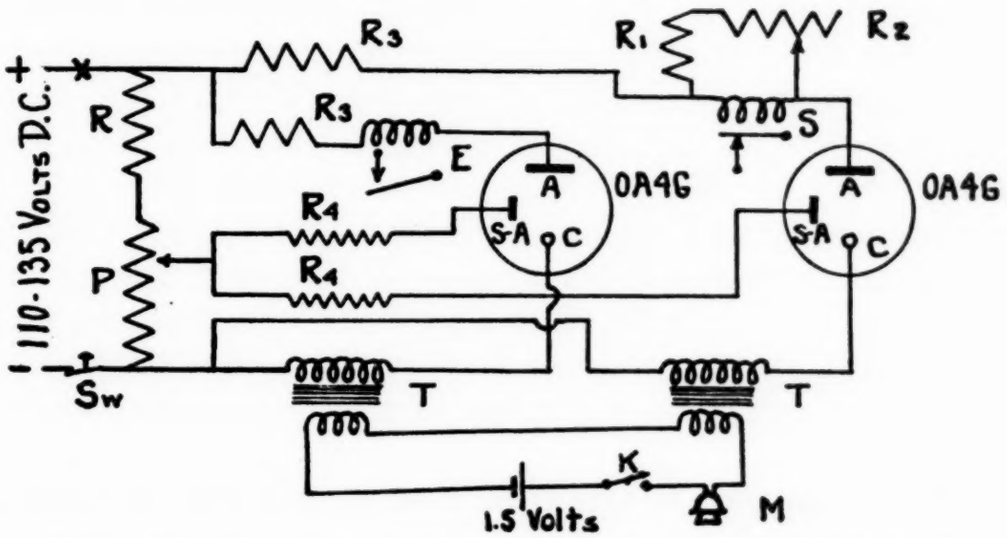

FIG. 2. Transformer primaries connected for adjustment of time of relay *S*.

enable the operator to make an additional small adjustment in the time delay of the *S* relay, without recourse to the relay armature itself, the resistors R_1 and R_2 are shunted across the *S* relay and variation of R_2 permits slight but sufficient adjustment of the time delay of the *S* relay so that this delay may be checked and adjusted if necessary from time to time.

For satisfactory operation it is best to so adjust the relays that a definite time delay is inherent in the *S* relay. This delay is readily measured by the chronoscope itself using the simple circuit of Fig. 2. This circuit permits actuation of both the *E* and *S* simultaneously by closure of switch *K*. Using this circuit the relays and R_2 are adjusted so that for simultaneous operation of *E* and *S* a small but measurable delay results. This measured delay is readily subtracted from

the measured reaction times. This procedure accounts accurately for any inherent differences in the operation of the two circuits.

Actually the adjustment can be carried to the point where no time delay occurs and at this point it would be unnecessary to correct the subsequent measurements. On the other hand, if a small delay is permitted one is assured that the relays and circuits are functioning properly; this certainly seems to warrant the above suggested method of procedure since it assures that the relay contacts are operating in the proper sequence.

In the circuit of Fig. 1 the direct current source of supply is shown as 110–135 volts. The relay will operate satisfactorily for any value of voltage between these limits. The usual direct current supply available in the laboratory is in this range and hence the voice relay can be operated directly from the laboratory line. Since the circuit will operate entirely satisfactorily at 135 volts; where a laboratory direct current supply is not available three 45 volt '*B*' batteries may be used. Note that switch *Sw* is provided to de-energize the circuit and extinguish the tubes after each operation of the relay.

The advantages of this vocal reaction key over previously described devices are:

(1) Cost: the entire apparatus necessary to build the voice relay can be purchased for approximately $25.
(2) Flexibility: only a direct current source of 110–135 volts is required; no alternating current is necessary.
(3) Accuracy: by the simple procedure described above one can account for time delay in the voice relay circuit.
(4) Action: action is positive and simple, no critical adjustments are necessary and no unduly loud sounds are required for operation.

(Manuscript received February 27, 1939)

References

1. Dorcus, R. M. and Hamburger, F., Jr., An inexpensive electronic chronoscope, *J. Gen. Psychol.*, 1938, **18**, 439–445.
2. De Silva, H. R., A timer for vocal reactions with thyratron tubes and microphones, *Amer. J. Psychol.*, 1935, **47**, 315–316.